The Spatial Ecology of Host Parasite Communities

Thesis submitted in accordance with the requirements of the University of Liverpool for the degree of Doctor in Philosophy by Shaun Patrick Keegan

August 2019

TABLE OF CONTENTS

ACKNOWLEDGMENTS

I would first like to thank Prof Andy Fenton for his supervision these past 4 years. On top of the already high intensity of PhD life, circumstances have proven difficult personally these last few years, and Andy's compassion and patience have proven every bit as invaluable as his role as my mentor. The independence I have been afforded in my PhD to explore new ideas, to strike up collaborations and to travel the world to discuss and improve my science has made me a better scientist is a testament to Andy's excellent supervision.

I would also like to extend my thanks to my wider supervisory team, Dr Amy Pedersen and Prof Mike Begon, for their input and counsel in how best to proceed and develop the research topics at hand. Additionally, to Dr Vanessa Ezenwa for allowing me access to a rich dataset of the African buffalo and supporting my short-term fellowship at the University of Georgia – it was an excellent experience.

I must also pay thanks to the many field and laboratory workers over the years who I do not know, but without which I would not have had the phenomenal datasets that I have with which to explore my ideas and interests in spatial disease ecology.

There are too many office mates and fellow PhD students to name all individually, but I would like to thank Kayleigh Gallagher and Toby Irving in particular, for being available to let me articulate my ideas, talk through debugging my code and generally being a sounding board for venting PhD related stresses to.

I would also like to thank Amy Halliday, David Webb and Steven Anderson, great friends who have made sure that I have never felt far from home while undertaking my PhD.

To my parents Jacqueline and Stephen, for supporting me and encouraging me to achieve my best before, throughout and after my PhD. I would not have made it through without their support.

Finally, I would like to thank my two adorable balls of fluff, Daisy and Riley, for never failing to give me a reason to smile, even when times are tough.

ABSTRACT

Tobler's First Law of Geography states that '*everything is related to everything else, but near things are more related than distant things'*. In the context of infectious diseases this implies that there are likely to be further cases of these infections spatially closer to an infected host than spatially further away. This is a simplistic interpretation that fails to consider the transmission biology of parasites. I investigated whether a function of spatial clustering, the K function, differed from the host and between parasites of a range of taxa and transmission modes in the wild wood mouse. Using the studentized permutation test, I found that there is a significant difference between a close contact transmitted virus, Wood Mouse Herpes Virus (WMHV), and the host and a number of parasite species. This is consistent with prior studies of close contact transmitted viruses in wood mice, suggesting there is a link between spatial clustering and close contact transmission.

Interactions among coinfecting parasites are typically examined at the within-host level, often revealing strong effects on individual host susceptibility or disease progression. However the effect of these interactions on parasite transmission between hosts, and the spatial scales over which those effects operate, has remained unknown. I analyse a spatially explicit dataset of the diverse community of parasites infecting wild wood mice, to assess the effects of local neighbourhood prevalence of each parasite species on individual-level infection risk by the other species, over an increasing range of spatial scales. This revealed that the effects of within-host interactions between coinfecting parasites can indeed ripple out beyond the individual host, resulting in a network of facilitatory and suppressive effects on transmission among these parasites. However

these between-host effects were only seen over relatively restricted distances around each host, over spatial scales likely reflecting the spatial scale of transmission.

Classical models of infectious diseases generally assume random mixing of individuals in a population. In these models each individual is as likely to encounter every other individual equally. Ignoring heterogeneities in contacts between individuals can overlook a significant element of the transmission biology of the parasite. Recently, studies focusing on how social networks relate to the spread of infection have increased in number dramatically. I explore how two measures of an individual's place in a social network, eigenvector centrality and degree, affect the disease status of individuals, using a very different study system to the wood mouse, the African buffalo. I find that for some parasites, eigenvector centrality affects disease status but that for all parasites degree has no effect. I then adapt the neighbourhood analysis technique to investigate potential novel parasite-parasite interactions, detecting one previously unknown.

Spatial scale is the theme binding each of the studies in this thesis - from scale of clustering, to scale of coinfection interactions. Using pre-existing and bespoke techniques, I have explored 2 very different host-parasite communities to tackle these issues, concluding that spatial scale is an important consideration in understanding parasite biology.

1 INTRODUCTION & LITERATURE REVIEW

1.1 Why space is important for epidemiology?

Infection can be characterised by the point in time that it occurs, and the point in space that it occurs (Real and Biek, 2007). However, traditional approaches for understanding and modelling the transmission and spread of infectious diseases adopt a 'mass action' approach that assumes homogenous mixing such that all individuals are equally likely to contact all other individuals in the population (Anderson and May, 1992). As such this standard approach ignores, or at least averages over, the spatial context of transmission. In reality, individuals may be more likely to contact individuals that are close to them, or there may be environmental heterogeneities that mean the standard assumptions of homogenous transmission across the population are invalid. As such, recognising the spatial and social context of transmission – with who and where contacts are most likely to occur – can dramatically improve our understanding of disease spread, and the development of more effective, targeted mitigation strategies. As an example of this, the most heavily cited (788 at the time of writing) article exploring the dynamics of the 2001 foot and mouth disease (FMD) epidemic in the United Kingdom was a spatially explicit, individual based model (Keeling *et al.*, 2001). While this approach is not always viable, the degree of information about the occurrence of cases on farms in the United Kingdom allowed for a suitably detailed model of transmission throughout the epidemic, with reference to the spatial arrangement of infected farms to one and other. This spatially explicit modelling allowed for clear recommendations to be provided on how to manage a subsequent FMD outbreak in the United Kingdom (Woolhouse, 2003). Traditional modelling approaches, that typically assume a form of homogenous

mixing (Anderson and May, 1992) would have failed to capture the dynamics of the spread of the virus.

Spatially, there are several types of parasite spread (White and Forester, 2018). Firstly, parasites can spread from an epicentre, uniformly, such as the case of West Nile Virus (LaDeau *et al.*, 2008). Some spread via dispersal events, over both small and large distances, depending on landscape features acting as barriers to dispersal or not, such as in rabies (Smith *et al.*, 2013). Others spredominantly pread locally, with very occasional long range dispersal events, such as the amphibian disease ranavirus (Price *et al.*, 2016). That there is such diversity in spread suggests that there is some element of the underlying biology, such as host or vector movement, or the nature of environmental dispersal, that dictates why one parasite spreads around an epicentre uniformly, and one will have large-scale dispersal.

1.2 Spatial scale

Scale is a popular, if problematic word in ecology. Nonetheless, it is a key concept of importance in ecological theory (Levin, 1992). It has many, often contradictory usages including the spatial and/or temporal resolution of a process, the extent (area considered by the study) of processes (Dungan *et al.*, 2002), or the level of biological organisation being considered (individual, population, community etc). In this thesis I specifically use the term to relate to spatial scale (rather than, for example, biological scale), but even then this is another concept with a somewhat over-stretched definition in the literature. There are three spatial patterns common to ecological data that affect observed processes. Firstly, at the largest of spatial scales, there are overall trends such as climate which is the dominant process. Secondly, at intermediate spatial scales there is patchiness of the environment, which is a dominant feature. And finally, at the

smallest of spatial scales there is stochasticity, whereby local variability dominates (Fortin and Dale, 2005). For the purposes of parasite ecology in the context of this thesis, I am interested in the 2 smaller scales, the patchy heterogeneous environment and the individual variability.

1.3 Transmission mode and the spatial clustering of infections

Tobler's First Law of Geography states that '*everything is related to everything else, but near things are more related than distant things'* (Tobler, 1970). In the context of parasites, this implies that susceptible individuals closer to an infected host would be more likely to be subsequently infected by that host, than individuals further away. If individual *a* is infected, and individual *b* is closer to *a* than individual *c*, the intuitive assumption is that individual *b* is more likely to become infected than individual *c*. This is an attractive view if one is considering only close contact transmission, however given the range of methods by which parasites infect susceptible hosts (Antonovics *et al.*, 2017), it may be an over simplification. If the parasite in question is transmitted by an arthropod vector, this may facilitate longer range transmission, particularly if that vector is a strong flying species such as a mosquito or tsetse fly. As such, individual *c* may be equally or even more likely to become infected than individual *b*. In this case, parasite transmission is dependent on the movement and distribution of the vector as well as the host movement and distribution. Other parasites transmit through shedding infective particles into the environment. Transmission is also potentially greatly variable among these parasites, as some like cholera (*Vibrio cholerae*) are water borne, and as such dispersal and transmission may be long-range, carried by flowing water, and spatial clustering may be apparent around sources of water specifically. Others, such as soil transmitted nematodes have limited dispersal capability in the environment, and so cases may be clustered where hosts occupy space in the environment.

To the best of my knowledge, no comparative study of parasite clustering between species with different transmission modes has ever been undertaken. A number of individual parasites have been examined (Carslake *et al.*, 2005, Ngowi *et al.*, 2010, Yohannan *et al.*, 2014), as well as comparing a single parasite between different hosts (Carslake *et al.*, 2005). It is this variability between transmission modes, and the implications it has for the degree to which cases of infection by the different parasites cluster around one and other; that motivates the work presented in this thesis.

1.4 Quantifying spatial clustering

A spatial point pattern is a dataset in which the locations of observations (as x and y coordinates) are recorded within a defined area (Diggle, 2013; Baddeley, Rubak and Turner, 2015). There is a common measure of spatial clustering, the K function (Ripley, 1976), that calculates the correlation between points in a spatial point pattern. This is usually expressed as K(r), indicating the clustering occurring over spatial distance r, whereby high values of K(r) indicate more clustering than would be expected than random. Spatial K functions have been employed to assess spatial clustering in a range of parasite systems, from the bacterium *Chlamydia trachomatis* (Yohannan *et al.*, 2014) to the zoonotic tapeworm *Taenia solium* (Ngowi *et al.*, 2010). A number of studies (Ribeiro *et al.*, 2015; Leite Dias *et al.*, 2016; Madinga *et al.*, 2017) employ a transformation of the K function, the L function. Any clustering detected that is above what is expected under complete spatial randomness is the product of processes which occur at spatial scales up to the point that clustering is detected, not simply at that spatial scale itself. There are yet further extensions of the K function, including the spatiotemporal K function (Diggle, 2013). These have been used to investigate clustering of cowpox virus in wood mice and bank voles (Carslake *et al.*, 2005) and FMD in cattle (Picado *et al.*,

2011). Spatially the function was used to determine that clustering of cowpox cases was highest within one home range radius of hosts (Carslake *et al.*, 2005). However, while this spatiotemporal adaptation of the K function is useful for the analysis of parasite clustering, given data of sufficient spatial and temporal resolution, it will not be used in this thesis for reasons outlined in further detail in Chapter 3.

1.5 The application of network theory to epidemiology

Classical models of infectious diseases generally assume random individuals mix randomly in a population, much like a chemical reaction. In these models each individual is as likely to encounter every other individual as much as any other individual (Anderson and May, 1992). In some cases, these models adequately describe the dynamics of the infection across the host population, it would be flippant to ignore heterogeneities in contacts between individuals which can overlook a significant element of the transmission biology of the pathogen - driven by variation between individuals in host behaviour. In the last decade or so, studies focusing on how host social networks relate to the spread of infectious diseases have increased dramatically (White, Forester and Craft, 2017). Social networks are representations of potential pathways for contact, and disease transmission to occur (White, Forester and Craft, 2017). The aim of utilising social network representations of populations is not to discount the utility of traditional compartmental models. Instead it is rather to understand fully the importance of heterogeneities in contact structure and understand how these contribute to the dynamics of disease in the system. Network analyses have been employed in empirical and modelling studies of infectious disease transmission across a wide variety of host-parasite systems, ranging from reptiles (Godfrey *et al.*, 2009; Aiello *et al.*, 2014), ungulates (VanderWaal *et al.*, 2014) to marsupials (Corner, Pfeiffer and Morris, 2003), primates (Griffin and Nunn, 2012; Carne *et al.*, 2014;

Romano *et al.*, 2016) to humans (Bansal, Grenfell and Meyers, 2007; Salathé and Jones, 2015).

It is traditional ecological thinking that group living should bring with it increased parasite transmission as a consequence (Altizer *et al.*, 2003). However, it has been demonstrated via modelling studies that social structures can lead to a protective effect from infection (Hock and Fefferman, 2012). The European Barger (*Meles meles*) lives in structured social groups and are carriers of bTB. A government program of badger culling has been undertaken to reduce transmission of bTB to cattle, causing disruption to the badger clans. This failed to consider the underlying social structure of badgers. Perturbation of these clans led to an increase in inter-clan interactions for up to 8 years after culling, and as a consequence increased infection (Carter *et al.*, 2007). Small-scale culling efforts are likely to result in increased bTB in cattle (Bielby *et al.*, 2014). This example showcases the importance of network structures and how they can seriously impact disease dynamics under perturbation.

1.6 Overview of study systems

Data from two very different study systems are analysed in this thesis, described in more detail in Chapter 2. Firstly, a longitudinal and spatially hierarchical mark-recapture dataset of wood mice and their parasites. Data on parasites in this system range from helminth worms and coccidial gut parasites to blood-borne bacterial and protozoan parasites. They represent a range of transmission modes, including close contact, environmental transmission and transmission via a vector. This dataset has been explored in a number of publications (e.g. Knowles *et al.*, 2013; Withenshaw *et al.*, 2016), however, until now has not been explored in a spatially explicit fashion. The diversity of

parasites recorded in this spatially explicit data set facilitate analysis of how the spatial clustering of cases varies across their differing transmission modes. The second system to be investigated is the African buffalo. Again, there is parasitological data for a range of micro and macroparasites. However, instead of spatially explicit data, the buffalo data is more implicit in the form of association data. This allows for the building of social networks, that can then be analysed to understand how social proximity affects infection occurrence for the range of parasites recorded.

1.7 Coinfection

As well as it being a truism that individuals in natural populations do not mix homogenously, another facet of natural infectious disease systems is coinfection; the simultaneous infection of individual hosts by multiple parasite species (Petney and Andrews, 1998; Cox, 2001). There is a great deal of ecology operating among parasites within hosts (Pedersen and Fenton, 2007) and there are a number of mechanisms by which parasites can interact. Such interactions have been observed in a number of wild host-parasite systems. For example, in the wood mouse (*Apodomus sylvaticus*), there has been a well characterised, antagonistic, interaction between a nematode, *Heligmosomoides polygyrus* and a coccidia, *Eimeria hungaryensis* (Knowles *et al.*, 2013). This was determined by experimental suppression of *H. polygyrus* via treatment with the anthelmintic Ivermectin, which resulted in a 15-fold increase in coinfecting *Eimeria* following nematode suppression by the drug. Conversely, in the African buffalo (*Syncerus caffer*) a positive association was detected between coinfecting strongyle nematodes and coccidial parasites (Gorsich, Ezenwa and Jolles, 2014), as well as a host of other interactions, including interactions between *Mycobacterium bovis* (bTB) and strongyle nematodes (Jolles *et al.*, 2008), and between bTB and Brucellosis (Gorsich *et al.*, 2018a). This latter interaction highlights the importance of scale, as mentioned above. bTB is a

risk factor for acquiring Brucellosis at the individual level, whereas at the population level, Brucellosis has a negative association with bTB (Gorsich *et al.*, 2018b).

1.8 Thesis outline

The overarching aim of this thesis is to understand the spatial and social context of infectious disease occurrence across the range of parasites recorded in the study systems, with the intention of using these systems to inform more generally how differing transmission modes affect the spatial spread and occurrence of infections. In Chapter 2 I describe in more detail the datasets that I will use. This will encompass the wood mouse mark recapture dataset, to be used in Chapters 3 and 4, and the African buffalo social network dataset to be used in Chapter 5. Chapter 3 explores the spatial clustering of parasite infection, with the specific focus being what scale parasites show clustering at, in a comparative way between parasite species. Chapter 4 focuses on coinfection, and present a novel technique to determine the spatial scale over which the effects of interactions between coinfecting parasites spread beyond the individual host. This uses the above well characterised interaction between *H. polygyrus* and *E. hungaryensis* (Knowles *et al.*, 2013) to validate the technique, which is then applied to other pairs of coinfecting parasites. Chapter 5 investigates social networks of the African buffalo and determines what individual level network characteristics drive infection. Additionally, in Chapter 5 I adapt the neighbourhood technique developed in Chapter 4 for use on social networks, to determine how parasite infection status is affected by prevalence among their neighbours at various distances in the network, and to similarly use this technique to investigate the scaling of coinfection interactions in this system. Chapter 6 presents an overall Discussion of the work presented in this thesis, and considers the broader applications of the findings presented for our understanding of the spatial spread and control of infectious diseases more generally.

2 SPATIAL CLUSTERING OF PARASITES WITH DIFFERING TRANSMISSION MODES

2.1 Abstract

Tobler's First Law of Geography states that '*everything is related to everything else, but near things are more related than distant things'*. In the context of infectious diseases this implies that there are likely to be further cases of these infections spatially closer to an infected host than spatially further away. This is a simplistic interpretation that fails to consider the transmission biology of parasites. Here I investigated whether a function of spatial clustering, the K function, differed from the host and between parasites of a range of taxa (virus, nematoda, bacteria and protozoa) and transmission modes (close contact, environmentally transmitted and flea borne) in the wild wood mouse. Using the studentized permutation test, I found that there is no significant difference between the host and environmental or flea borne parasites, but that there was a significant difference between a close contact transmitted virus, Wood Mouse Herpes Virus (WMHV), and the host and a number of parasite species. WMHV clustered significantly more at spatial scales of up to ~25 metres. This is consistent with prior studies of close contact transmitted viruses in wood mice, suggesting there is a link between spatial clustering and close contact transmission.

2.2 Introduction

As early as 1854, in his now famous disease maps, John Snow recognised there was spatial clustering of cholera cases in London (Snow, 1855). This led to the identification of the Broad Street water pump as the most likely source of disease; cases became less frequent the further away from the Broad Street pump they were, or the closer to another water pump. Snow's work was a key moment for epidemiology as it showed that by using spatially-explicit disease incidence data it is possible to track down the source of infection, and its conclusions have been verified and built upon by numerous studies. For example, using Kernel Density Estimation, Shiode *et al.* (2015) showed quantitatively that the greatest mortality from cholera was found to be clustered spatially around the Broad Street pump. Snow's cholera data have also been used to develop geographic profiling tools for epidemiology, using what we know about the clustering in that situation to determine whether new approaches are also able to identify the Broad Street pump as the source (Papini and Santosuosso, 2017).

The clustering of cholera cases that Snow observed is in line with Tobler's First Law of Geography, that '*everything is related to everything else, but near things are more related than distant things'* (Tobler, 1970). In the context of infectious diseases, this makes the intuitive assumption that susceptible individuals closer to an infected host would be more likely to be subsequently infected by that host, than individuals further away. While an attractive generalisation, this may be a reductive view. The spatial scale of clustering is likely to depend on a number of factors related to both the host, the parasite and the environment. The fundamental process required for the propagation of disease is transmission between host individuals. Parasites (defined here as an organism living in or on another, and that causes harm to that organism), or their infectious propagules

must pass from one host to another. There are a variety of methods that parasites use to transmit from host to host, some passive (e.g. *Ascaris lumbricoides* eggs ingested from the environment) and some active (e.g. schistosome cercariae seeking a host to infect in water), some short-range (e.g. sexually transmitted parasites) and some long-range (e.g. tsetse fly transmitted trypanosomiasis), leading to potentially very different patterns of spatial clustering, over very different spatial scales. Cholera is water borne, which is why there was clustering of cases centred on a contaminated water pump, at the smallest spatial scales. Similarly, for parasites that transmit by close contact between hosts such as influenza, we may expect cases of infection to cluster closely at small, local spatial scales. However, other modes of transmission may give rise to very different clustering patterns. For example, for parasites that transmit by using an arthropod vector, transmission is dependent not only on host movement and distribution, but on the movement and distribution of the vector. An example of this is malaria, transmitted by the mosquito *Anopheles gambiae*. The mosquito is able to move between hosts independently, and this extensive vector movement, or prolonged survival in the environment, may decouple the clustering of cases from that of the host, and we may not expect to see localised clustering of cases. In the case of environmentally transmitted parasites, whether the substrate (soil or water) is static or mobile (in the case of flowing rivers), and whether it is ubiquitous, such as soil, or forms discrete patches in the environment, will again influence the spatial scale of any clustering of cases. For example, if the medium for transmission is the soil, we may see clustering that is closely related to the territoriality of hosts. Therefore, the scale at which we observe clustering of infection in relation to host clustering, can provide biological insights into the host-parasite dynamic, specifically the spatial scale of transmission.

2.2.1 Studies into the clustering of infectious diseases

A number of studies have explored aspects of spatial clustering for a variety of infectious diseases. For example, in a large study of HIV in rural Uganda, Grabowski *et al.*, (2014) investigated spatial clustering of cases across a variety of spatial scales, from within household (0 metres apart) up to 250 kilometres. Individuals within households were 3.2 times more likely to be seropositive for HIV than the general population examined in the study. Weaker, although still statistically significant, clustering was detected beyond the household, with individuals within 10 – 250 metres of each other being 1.22 times more likely to be seropositive than other participants in general, and 1.08 times more likely when occurring between 250 – 500 metres of each other (Grabowski *et al.*, 2014). HIV is transmitted by sexual (close) contact between individuals, limiting its ability to spread over large scales spatially by the constraints on its host. The evidence presented by Grabowski *et al.*, (2014) demonstrates this in the form of strong local clustering, with weaker clustering as spatial scale increases.

Influenza is an orthomyxovirus, responsible for the deaths of between 50 and 100 million people in the first pandemic outbreak of the 20th century alone (Johnson and Mueller, 2002). Seasonal epidemics occur, due to small changes to the viral surface proteins (Smith *et al.*, 2004), whereas pandemics occur when viruses present novel surface antigens (Simonsen, 1999). Patterns of host movement, as opposed to population size or density were found to be the key determinant of spread of infection of interpandemic, seasonal influenza in the United States of America (Viiboud *et al.*, 2006). Hence the disease spreads in a spatial manner consistent with the movement of the hosts.

2.2.2 Quantifying clustering: spatial point patterns and K function analysis

When considering spatial clustering, it is important to think about the type of data required for a specific analysis. A spatial point pattern is a dataset in which the locations of observations (as x and y coordinates) known as *events* are recorded within a defined area, or *window* (Diggle, 2013; Baddeley, Rubak and Turner, 2015). Ripley's K function is a standard tool used to determine over what scales a spatial point pattern clusters (Ripley, 1976). This function *"describes the extent to which there is a spatial dependence in the arrangement of events"* (Gatrell *et al.*, 1996) where, in the context of infectious diseases, an event can be a parasitic infection. While in many ecological studies of K functions, the aim is to determine whether there is greater clustering than one would expect by chance (i.e., if events were distributed randomly and independently of each other in space), in the case of parasitic infections, an important question is whether a parasite shows greater or lesser clustering than the host. Should a parasite be more clustered than the host population, it would suggest that there is spatially localised transmission, as opposed to the mass action type of transmission that forms the basis of most of our understanding of infectious disease ecology (e.g., Anderson and May (1992)). As an example of this kind of analysis, Carslake *et al.*, (2005) hypothesised that for cowpox, a virus of rodents in the UK transmitted by close contact, cases would be spatially clustered within the home range of the host, and temporally clustered within the time it takes for the host to rid itself of the virus (approximately 4 weeks). Developing an adapted version of the space-time K function, for both host species (wood mice and bank voles), Carslake *et al.* (2005) found that clustering of cowpox cases was indeed constrained to a single home range of the host (~ 16 metres) and temporally clustered within 4 weeks. This is biologically important given the main route of transmission of cowpox is via close contact between individuals. Close contact transmission requires that both infectious and susceptible animals be at the same point in space at the same

point in time; hence the observed spatiotemporal scale of clustering of cases match that expected based on the transmission biology of the parasite. That is, that it would be found within the range of space occupied by a single host, within the infectious period of the virus. Like HIV, influenza, and measles, this evidence suggests that clustering for parasites transmitted via close contact should be driven by host movement and distribution, and as such is likely to be highest at the scale of space use by the host. This approach proved useful in a subsequent study, whereby the spacetime clustering of mice and voles individually fell within the expected spatial and temporal scales for each species alone, but did not cluster with respect to the distribution of the opposite species, suggesting little or no cross-species transmission (Carslake *et al.*, 2006). Hence the relationship between the clustering of disease cases and the host can provide invaluable information about potential routes and modes of transmission.

2.2.3 A natural model of clustering of diverse parasite types: The wood mouse parasite system

The wild wood mouse (*Apodemus sylvaticus*) is native to the United Kingdom, as well as much of Western Europe. These animals are host to over 30 species of parasites, of varying taxa and transmission modes (Knowles *et al.*, 2013). This diversity across the parasite community makes these animals an ideal study system to investigate clustering of infection cases in relation to that of the host, and between the different parasite species. As described above, previous clustering analyses have been undertaken in the wood mouse and bank vole (*Clethrionomys* [now *Myodes*] *glareolus*) system, with regards to the host populations and of individuals infected with cowpox virus (Carslake *et al.*, 2005). However, cowpox is just one, relatively rare (~5%) parasite in this system; there

are a large number of parasites showing a range of transmission modes and chronicities (Table 2.1).

Table 2.1 Parasites considered in the clustering analysis, their transmission mode and chronicity of infection.

Parasite	Taxa	Infection Type	Transmission Mode
Heligmosomoides polygyrus	Nematoda	Chronic	Environmental
Eimeria hungaryensis	Protozoa	Acute	Environmental
E. apionodes	Protozoa	Acute	Environmental
Bartonella grahammi	Bacteria	Acute	Vector (Flea)
B. taylorii	Bacteria	Acute	Vector (Flea)
Trypanosoma grosi	Protozoa	Chronic	Vector (Flea)
Wood Mouse Herpes Virus	Virus	Chronic	Close Contact or Vector (Tick)

The most common parasite in the system is the nematode *Heligmosomoides polygyrus*, which is transmitted environmentally in the soil. While Snow's cholera had a clear environmental source to cluster around (water pumps), given that the nematode eggs are excreted as the mice defacate as they move across the environment, clustering for this parasite may be less local. This may also be true for the similarly transmitted protozoan parasites *Eimeria hungaryensis* and *E. apionodes*, which infect faeco-orally via occysts in the envionment. However, given that they have a differing chronicity of infection compared to *H. polygyrus*, being acute rather than chronic, this may yet again affect the scale over which cases of infection cluster. In particular, we may expect a chronic environmentally transmitted parasite to be continually shedding infectious particles into the environment, whereas for acute infections, there is a much more restricted temporal window by which these particles can be shed. Hence it is reasonable

to hypothesise that *H. polygyrus* will be clustered closely with the host in space, whereas for *Eimeria spp.* there may be a temporal decoupling, which may cause it to be less consistent with the host with regards to spatial clustering. Contrastingly, vector transmitted parasites may show different patterns of clutering compared to parasites with these other transmission modes. Many vectors can move independently of their hosts, and as such two hosts do not need to come into contact spatially or temporally for transmission to occur. Fleas, for example can potentially move far from their host relative to their size, and may increase the range spread, particularly spatially. Tick borne parasites may be likely to show significant temporal decoupling from the host, given ticks consume a blood meal, leave that host and then do not consume their next blood meal until the following year. Hence, the clustering of tick-borne pathogen cases may be quite unnconnected to the contemporary clustering of their hosts. The wood mouse is infected by flea borne parasites such as the chronic-infecting protozoan *Trypanosoma grosi*, and multiple species of the acute-infecting bacteria *Bartonella spp.*. Hence, comparing between these different parasites allows me to ask whether the degree of clustering of infections is the same for parasites with the same tranmission mode (indeed, the same vector species), or if there is some other aspect (e.g. chronic vs. acute infection) that differentiates how cases of parasite infection are clustered. Finally among the parasites in this system that I consider is wood mouse herpes virus (WMHV). While it is commonly believed WMHV virus is transmitted via close contact, there is some debate as to its true route of transmission in the field, and there is some evidence to suggest that it is also transmitted by tick vectors (Hajnická *et al.*, 2017). Hence it is possible that the scale over which WMHV infections cluster may give biological insight into which hypothesised transmission mode (direct contact v. tick-borne) contributes most to the overall transmission, and spatial distribution, of WMHV cases.

This chapter sets out to address two key questions relating to the spatial clustering of parasite cases in the wood mouse system. Firstly, are parasite cases more or equally spatially clustered than that of the host animal? Next, do parasites cluster differently from each other and, in particular, does any difference in the spatial scale of clustering correlate with transmission mode? To accomplish this, I will use a longitudinal dataset of wood mouse captures and parasitological data for these captures. Note that, whereas Carslake *et al.*, (2005) were able to look for temporal and spatiotemporal clustering, due to the long-term nature of their data, our data are temporally limited to monthly from May to December, which would severely restrict power of such analysis. Hence, due to this lack of temporal resolution and range, I have elected to focus exclusively on spatial clustering.

2.2.4 Hypotheses

Given the above information regarding the potential differences arising from transmission mode, I hypothesise that I will detect differences in the spatial scale of clustering between parasites of differing transmission modes. Additionally, I expect to see different clustering between parasites and their host, depending on the transmission mode, again with parasite species that share a transmission mode showing similarly-distinct clustering from that of the host.

2.3 Methods

2.3.1 Data

These data were collected under the supervision of Professor Andy Fenton and Dr Amy Pedersen over a 6 year period from 2009 to 2014. A number of papers have been published using parts of these data over the years (e.g. Knowles, Fenton and Pedersen,

2012; Knowles *et al.*, 2013; Withenshaw *et al.*, 2016), however all analysis in this thesis is novel.

Individuals were trapped three-weekly or four-weekly, using baited Sherman traps (Alana Ecology, UK; dimensions 8.9cm x 7.6cm x 22.9cm), from May to December in 2009 to 2014 in woodlands in the North West of England. Traps were laid on a semi-permanent grid with 2 traps laid every 10 m, which were checked every day for 4 consecutive days in trapping weeks. At first capture, all mice were permanently tagged with a tagged with a subcutaneous microchip transponder for identification (AVID Friend Chip). For all mice at each capture, the following metrics were taken: body length (nose tip to base of tail), weight (g), sex, reproductive status and an estimate of age were recorded (see Knowles *et al.*, (2013) for further details). At every capture, faecal samples were collected from previously sterilised, single occupancy traps for faecal floatation and microscopic analysis to identify and quantify both helminth eggs and coccidial oocysts (measured as eggs/oocysts/gram; see Knowles *et al.* (2013) for details of identification and quantification of infection by these gastrointestinal parasites). At each weekly capture, a small blood sample was taken from the tip of the tail for analysis of microparasites (including wood mouse herpes virus, and blood-borne infections *Bartonella* spp. and *Trypanosmoa grosi*; see Knowles *et al.* (2012) and Withenshaw *et al.* (2016) for details of identification and quantification of these parasites). For microparasites, sensitivity analyses of detection methods were limited due to the small volume of blood able to be extracted from each animal, and the restrictions on repeated bleeding. For macroparasites, there is some variation between captures for individuals with regards to the number of parasites detected (e.g. eggs/oocysts per gram). Repeatability for infection status was relatively high (*Eimeria spp.* 78.7%, *H. polygyrus* 73.7%) (A. Pedersen, pers comm). Individuals were also checked visually for

ectoparasites (fleas and ticks), as described in Withenshaw *et al.* (2016). All mice were released at the point of capture after handling.

Table 2.2. Number of hosts and parasite prevalence across all grids for each year as based on the first capture of each individual. Parasite prevalence presented as the mean across all grids investigated in each year. *Hp* = *H. polygyrus*; *Eh* = *E. hungaryensis*; *Ea* = *E. apionodes*; *Bgr* = *B. grahamii*; *Bta* = *B. taylorii*; Tryps = *T. grosi*

Year	Host	*Hp*	*Eh*	*Ea*	*Bgr*	*Btay*	*Tryps*	WMHV
2009	521	0.304	0.323	0.199	0.203	0.267	0.081	0.142
2010	434	0.430	0.253	0.166	0.245	0.289	0.159	0.165
2011	725	0.262	0.255	0.254	0.131	0.175	0.067	0.094
2012	571	0.164	0.191	0.156	0.088	0.342	0.058	0.111
2013	471	0.085	0.101	0.165	0.031	0.159	0.070	-
2014	981	0.182	0.171	0.108	0	0	0.054	-

The sampled population represents approximately 80% of the total population of the grid (A. Fenton, pers. comm.). Given that individuals are often caught multiple times, the decision was taken to only consider the first capture of an individual in a given year, as a way to standardise what data was used in the final analysis, relative to that individual. While it was known to occur, individuals rarely crossed between grids, likely owing to the small home ranges used by wood mice.

2.3.2 K function analysis

The K function is a widely used tool for determining whether a point pattern is more or less clustered than would be expected under the conditions of complete spatial randomness (CSR). The formula of the K function over spatial distance *r* is:

$$\widehat{K}(r) = \frac{|W|}{n\,(n-1)} \sum_{i=1}^{n} \sum_{\substack{j=1 \\ j \neq i}}^{n} I\{d_{ij} < r\}\, e_{ij}(r) \qquad (3.1)$$

Where $|W|$ is the observational window (i.e., number of trapping points on the grid), n is the number of points in the point pattern (i.e., within distance r of each other), and d_{ij} is the distance between points i and j. $I\{d_{ij} < r\}$ is an indicator function, which takes the value 1 if $d_{ij} < r$ (i.e., if points i and j lie within distance r of each other, and hence should be included in the calculation of the K function), and 0 otherwise. $e_{ij}(r)$ is the edge correction (Baddeley, Rubak and Turner, 2015). The edge correction is included to account for the fact that points towards the edge of the window will have neighbours that are undetectable out-with the study window.

It should be noted that the data were collected from traps laid in pairs every 10 metres. This leads to the issue that there is some small scale variation in trap locations. Hence there may be two recorded captures of animals at a given trap location (one in each of the pair of traps at that location). Given the inability of the unmarked K function to deal with multiple replicates of the same point, the decision was taken to jitter the spatial coordinates by up to 1 metre, to account for small-scale, but unquantified, spatial variation in trap position around each location. This was preferred to the marked K function, which can allow for multiple entries per point location, as in the absence of fine resolution GPS data, it was considered the best way to consider multiple captures at the same trapping date and same coordinates using the 10 metre trapping regime.

All K functions and subsequent analyses were generated in R version 3.5.2 (R Core Team, 2018) using the package "spatstat" version 1.58-2 (Baddeley, Rubak and Turner, 2015).

2.3.3 Statistical analysis of K functions

Testing whether a K function differs from CSR is undertaken by enveloping of the K function. This is done by simulating point patterns to produce an outward bound from which, if the observed K function deviates, we can determine that it is significantly different from CSR (Baddeley, Rubak and Turner, 2015). Significant differences between groups of point patterns can be detected using the non-parametric Studentized Permutation Test (Hahn, 2012), which calculates a summary function, in this case the K function, for each pattern supplied to it, and then determines whether there is a statistically significant difference between these summary statistics between groups. The output is a p value which can be used to determine statistical significance. In what follows, each pattern represented either the host or cases of infection by one of the parasites for a given year on each grid. For example, cases of *H. polygyrus* on a single grid in a single year produced a point pattern, and this process was repeated for all parasites and the host for all years of the study. Due to the differing sampling efforts over the 2 survey periods (2009-2011 and 2012-2014), these periods were treated as independent groups of points and analysed separately. This was confirmed by running the Studentized Permutation Test on these two groups, which revealed significant ($p < 0.05$) differences in K functions of the host between the two year groups (2009-11 and 2012-14); however K functions within each year group did not differ significantly from each other.

Initially these statistical tests were run with point patterns of parasites and the host included to determine whether there was any significant difference between the degree of clustering of any parasite species and the host. This was carried out at a series of

spatial distances to assess whether, and over what spatial scales, host and parasite clustering differed. Firstly, this was done from 0-35 metres, to test whether there were significant differences in clustering over 2 home ranges around the host (Carslake *et al.*, 2005). Secondly, these tests were run on subdivided spatial scales, to assess any difference in clustering between parasite-parasite/parasite-host up to 1 home range (0-16 metres) and then from 1-2 home ranges (17-35 metres).

In cases where statistically significant differences were detected in the full models, these were then used to undertake Studentized Permutation Tests in a pairwise manner between all parasite-parasite/parasite-host pairs as a form of post-hoc analysis to identify which specific combinations (host and/or parasite) differed. In what follows we present both the raw p values from each pairwise comparison, and also when adjusted using the False Discovery Rate (Pike, 2011), reported as q in the results, which accounts for multiple testing and can be interpreted similarly (here, statistical significance is assumed for $q < 0.05$). As a supporting component of this post hoc analysis, K functions for each parasite/host were pooled, and then plotted for comparison.

2.4 Results

2.4.1 Differences between host and parasite clustering

For 2009-2011 there was no significant difference in clustering between the host and parasites at any spatial scale (0-35 metres: T =1316.4, p = 0.385, q = 0.385; 0-16 metres: T =227.9, p = 0.973, q = 0.973; 2012-2014: 17-35 metres: T = 1145.7, p = 0.078, q = 0.156). Similarly for 2012-2014 there was no significant difference in clustering between the host and parasites from 0-35 metres (T = 710.85, p = 0.039, q = 0.078) or from 0-

16 metres (T =115.81, p = 0.491, q = 0.654). However there was significant difference in clustering between the host and parasites from 17-35 metres (T =588.06, p = 0.007, q = 0.028). As a result, post hoc pairwise models were run using the Studentized Permutation Test (Table 2.3).

Table 2.3. p and q values of models testing for significant differences between K functions, run on the 2012-2014 data and over a spatial range of 17-35 metres. p values are shown in the top diagonal of the table and q values are shown on the bottom diagonal. *Hp* = *H. polygyrus*; *Eh* = *E. hungaryensis*; *Ea* = *E. apionodes*; *Bgr* = *B. grahamii*; *Bta* = *B. taylorii*; Tryps = *T. grosi*.

	Host	***Hp***	***Eh***	***Ea***	***Bgr***	***Bta***	**Tryps**	**WMHV**
Host	-	0.351	0.531	0.641	0.192	0.280	0.447	0.072
Hp	0.655	-	0.538	0.279	0.723	0.773	0.603	0.053
Eh	0.717	0.717	-	0.518	0.258	0.401	0.410	0.021
Ea	0.733	0.603	0.717	-	0.192	0.272	0.351	0.239
Bgr	0.603	0.774	0.603	0.603	-	0.648	0.774	0.024
Bta	0.603	0.774	0.675	0.603	0.733	-	0.655	0.029
Tryps	0.695	0.733	0.675	0.655	00.744	0.733	-	0.090
WMHV	0.4032	0.371	0.270	0.603	0.270	0.270	0.420	-

Based on a threshold for significance of $p < 0.05$, the majority of parasites did not have K functions that differed significantly from each other, or from the host (Table 2.3; Fig 2.1). The exceptions to this all involved WMHV; two parasites had significantly ($p < 0.05$) different K functions over 17-35 meters compared to WMHV: *B. taylorii* ($p = 0.029$; Fig 2.1) and *E. hungaryensis* ($p = 0.021$; Fig 2.1). In addition WMHV had a borderline significantly ($p < 0.1$) different K function from the host ($p = 0.072$; Fig 2.1), *H. polygyrus* ($p = 0.053$; Fig 2.1) and *T. grosi* ($p= 0.09$; Fig 2.1). In all cases WMHV displayed greater K(r) values (Fig 2.1), and so greater degrees of clustering, than these other parasite species, and the host, over 17-35 metres.

However, after p value correction with the false discovery rate, there were no models within the pairwise tests that were statistically significant ($q > 0.05$ for all comparisons; Table 2.3). This is not an unexpected issue when running pairwise models, given the stringency of the false discovery rate to ensure no false positives. To attempt to determine where the difference between the K functions of the point pattern groups lie, they were inspected visually (Figs 2.1). Overall it appears that the environmentally transmitted parasites *H. polygyrus*, *E. hungaryensis* and *E. apionodes* show patterns of clustering that closely follow those of the host (Fig 2.1), as do the two *Bartonella* species (Figs 2.1). However, WMHV appears to show higher degrees of spatial clustering than the host and the environmentally transmitted parasites *H. polygyrus*, *E. hungaryensis* and *E. apionodes* (Fig 2.1), and these differences are most strongly apparent at larger spatial scales (i.e., distances in excess of 20 metres). Furthermore, the flea-transmitted parasite *T. grosi* seems to show less clustering than the host and the environmentally transmitted parasites (Fig 2.1), and these differences are most apparent between 10-30 metres.

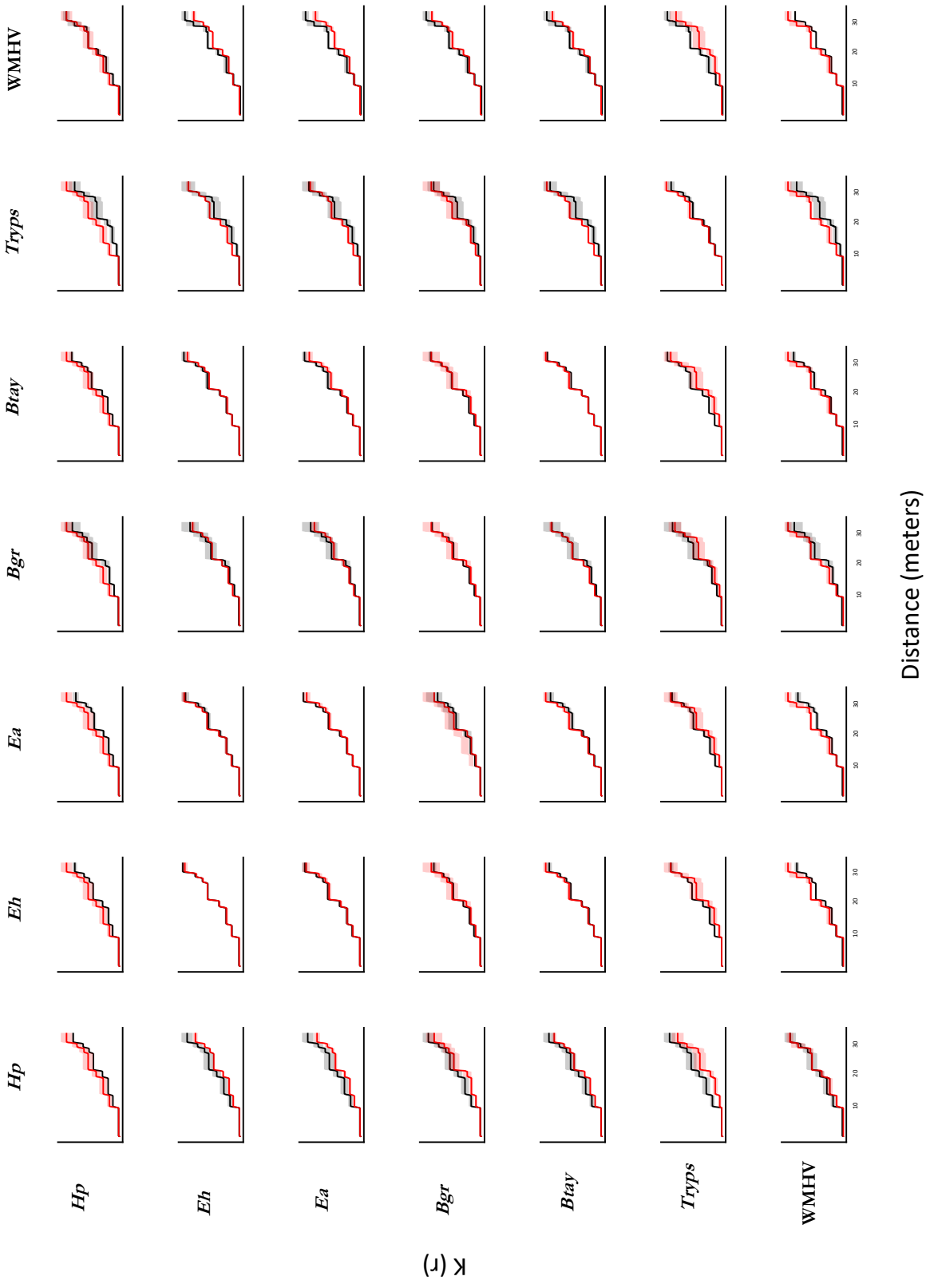

Figure 2.1 The spatial clustering (K function) of parasites in comparison with one and other over 0 to 35m. All data shown are from 2012-2014 combined. The parasite denoted on the left-hand side is shown in red, with the other parasite, denoted above, shown in black. The main diagonal shows clustering of the parasite defined on the left-hand side in red, and of the host in black.

An important point to consider when comparing K functions as in Fig 2.1 is that the lack of a difference in the K function does not equate to a lack of an interesting result. For example, consider the flea borne parasites. Here, a lack of difference is perhaps more interesting across all spatial scales given that fleas are able to move independently of the host.

2.4.2 Null Model Testing

A null model was used to check the approach was statistically robust and did not generate false positives. Data were simulated to conform with the prevalence of each parasite on each grid for the years 2009-2011. This was done by maintaining the spatial locations of the real captures of mice, and randomly assigning infection amongst them consistent with the prevalence of each parasite on that grid and year combination. The K function of these simulated data were compared to the real host K functions showing no significant differences (Table 2.4). This supports the analyses, in that a simulated parasite pattern does not diverge from the underlying pattern of host clustering.

Table 2.4. p values of models testing for significant differences between K functions of simulated infections with respect to the host in 2009-2011, across 0-35m, 0-16m and 17-35m

Parasite	p (0-35m)	p (0-16m)	p (17-35m)
H. polygyrus	0.906	0.386	0.200
E. hungaryensis	0.708	0.370	0.204
E. apionodes	0.922	0.366	0.196
B. grahamii	0.652	0.378	0.222
B. taylori	0.299	0.371	0.221
T. grosi	0.730	0.369	0.180
WMHV	0.911	0.372	0.187

2.5 Discussion

2.5.1 Apparently high degrees of clustering of Wood Mouse Herpes Virus

At scales greater than 20 metres, I found evidence that WMHV exhibits a higher degree of clustering than the host, as well as other parasites, both environmentally transmitted (*E. hungaryensis* and *H. polygyrus*) and flea-borne (*B. taylorii* and *T. grosi*). This is striking as WMHV is the only parasite I analysed to be transmitted via close contact of individuals. This increased clustering relative to that of its host, is similar to what Carslake *et al.* (2005) observed for another close contact transmitting virus, cowpox. While it is difficult to generalise too far beyond two examples, the similarities in these responses for these similarly transmitted parasites, do suggest contact transmitted viruses tend to show higher levels of spatial clustering than their host. Mechanistically this may arise because there is a need for both infectious and susceptible host to occupy the same point in space for the virus to transmit, therefore leading to highly localised transmission, rather than individuals at differing distances from each other being equally likely to be infected. However, it should be noted that Carslake *et al.* (2005) saw the highest levels of clustering of cowpox cases at particularly localised spatial scales, typically restricted to within one home range size of the host (~16 metres), whereas the clustering of WMHV cases observed here was seen at larger spatial scales up to two host home range diameters (35 metres). This may reflect genuine differences between their pathogens, in transmission mode or other aspects of life-history (see below for further consideration of these possibilities). However, it must be remembered that the K function is a cumulative measure. Whilst we may observe this increased relative clustering over 20 metres, processes operating at smaller spatial scales may be impacting this.

As noted in the introduction, it has been suggested with some experimental evidence (Hajnická *et al.*, 2017) that WMHV may transmit via a tick vector. If WMHV is transmitted by this route to a substantial degree, then this may explain why it exhibits clustering at spatial scales different from what Carslake *et al.* (2005) observed for cowpox. Furthermore, if tick transmission is important for WMHV, we may still expect it to differ from the other vector-borne parasites in this system (*Bartonella* spp. and *T. grosi*), which are flea transmitted. Fleas and ticks have substantially differing ecologies and host seeking behaviours which will impact their use of space. Fleas are relatively mobile, and able to move from host to host over some distance. Ticks on the other hand are significantly less mobile in space, and tend to seek one blood meal per year. This not only would result in differing spatial ecologies of transmission, but also differing spatiotemporal ecologies of transmission. So, while we may not see a difference between the K functions of WMHV and *B. taylorii* and *B. grahamii*, it is also not surprising that we do see a difference between the K function of WMHV and *T. grosi*, at scales above 20 metres. Unfortunately, with no data available on any tick-borne parasites in this system that are definitely transmitted by ticks, I cannot draw any firm conclusions about the transmission mode of WMHV via clustering comparisons, other than to note that it clusters in a distinct fashion from the host, some environmentally transmitted parasites, and some flea borne parasites.

WMHV also has another unique feature about its biology which may lead to a fuller understanding of parasite clustering. In our dataset, WMHV infections were quantified in terms of seropositivity, which detects host antibodies to WMHV, and so essentially measures whether the host has ever been infected by the parasite. For WMHV this is a rather crude and indirect measure as, although causing chronic infections, the parasite

has both an active and latent phase (Wu *et al.*, 2000). In the active phase, the parasite is able to transmit to other hosts, however in the latent phase it cannot. Hence where infected parasites are detected via serology may bear little resemblance to where they are actually transmitting to other hosts. Future studies should aim to resolve this issue, by searching for both seropositivity, to determine whether there is infection or not, but also for *Orf50*, a genetic marker of active WMHV infections, detectible by PCR (Wu *et al.*, 2000). Typically this is done using the spleen as the best site for detection from sacrificed animals (Hajnická *et al.*, 2017). However, should it be possible to detect this gene from smaller-volume blood samples also (i.e., without the need for destructive sampling of the host), this would provide a significant boost to understanding clustering of infection spatially. Doing so would not only produce an interesting temporal dataset of activation and latency of infection in the wild, but would allow us to determine whether active infections show spatial clustering, thereby indicating contemporary bouts of transmission.

2.5.2 Clustering of environmentally transmitted parasites

We found no evidence that environmentally transmitted parasites in our dataset (the nematode *H. polygyrus*, or the gut-dwelling *Eimeria* spp.) show significantly different clustering from the host animals over any spatial scales. This contrasts with what Snow found in the cholera cases in London, mentioned in the introduction, which is also environmentally transmitted. However, cholera was not ubiquitous in the environment, and the clustering was detected around sources of cholera (water pumps) as opposed to cases, potentially explaining the differences observed here. Wood mice defecate in the environment throughout their home range. This would imply that *H. polygyrus* eggs and *Eimeria spp.* oocysts are also deposited throughout their range. Given that *H. polygyrus* eggs and *Eimeria spp.* oocysts can remain viable for some time in the soil (likely several

months), the environment which it occupies is likely to pose an infection risk to others for a long period of time. Hence the deposition of parasite eggs or oocysts into the environment from an infected animal, and the subsequent uptake and infection of them into another animal will act to decouple the observed occurrence of cases. As such it is then not surprising that we see the distribution of parasites to be not different from that of the hosts.

2.5.3 Clustering of flea borne parasites

Perhaps most surprisingly in this study, *B. grahamii* and *B. taylorii* do not differ from host clustering at any spatial scale, whereas *T. grosi* appears less clustered than the host at ranges up to 25 meters in the 2012-2014 data, when assessing the K function plots visually (Fig 2.1). I had hypothesised that as fleas are able to disperse independently of the host, that I may see clustering operate at potentially larger scales than the host home range. However I found no support for this hypothesis. This leads to two potential conclusions. Either, flea vectors are spatially tied significantly closer to the host than I had anticipated, or the scale of analysis is insufficient to detect any difference in clustering. On the latter point, this study examines clustering up to 2 home ranges of the host (~35 metres). As the K function is a cumulative metric, it may be that clustering between 1 and 2 home ranges of the host only becomes apparent when looking over a larger spatial scale. A remedy to this in future studies, to ensure a fuller understanding of flea borne transmission is to use larger grid sizes. This would allow the K function to be calculated at larger scales, with limited edge effects. Currently, with the size of the grids, edge effects become increasingly important in the calculation of the K function, even for relative small spatial scales. A larger grid would reduce the importance of peripheral points and allow for larger spatial scales to be examined with confidence.

2.5.4 Lack of spatiotemporal analysis

While there has been no spatiotemporal analysis in this chapter, it is certainly the direction that future analyses of these types should take. Given this has been undertaken on a legacy dataset (i.e., one collected before the present analyses were conceived), I was unable to direct sampling with respect to a spatiotemporal analysis. The benefits to the spatiotemporal approach as undertaken by Carslake *et al.*, (2005) is that the coupling of cases in both space and time can be given due consideration. The key difference in the dataset they analysed and the one that I have analysed is that there is a significant gap in trapping each year (from January – May) in the data examined here, as opposed to full annual coverage across multiple years. This is the limiting factor in being able to combine temporal clustering with spatial clustering. Carslake *et al.*, (2005) were able to use their temporally extended data to not only determine that cowpox virus clusters spatially within one home range, but also that it clusters temporally within approximately 4 weeks of the infection being identified, consistent with what is known with the infectious period of cowpox virus. This may become important when trying to differentiate clustering among some of the parasites considered here. For example, while there was no difference seen in purely spatial clustering between *H. polygyrus*, *E. hungaryensis* and *E. apionodes*, this does not preclude the possibility that their clustering in time may differ (e.g., due to differences in longevity of environmental infection stages, for example), and as such their overall spatiotemporal clustering may be different. I would encourage future analyses on this type of system to consider this in full before sampling.

2.5.5 Sample Size and Inference

One point to consider when comparing clustering between different species of parasite and the host is the varying prevalence of each parasite. For the host-parasite comparisons, this represents comparing a K function to a subset of the same K function. Were prevalence to reach 100%, there would be no difference and we would be observing the same K function. This warrants caveating that for high prevalence parasites, this method may be insufficient for comparison. It is likely that the spatiotemporal methods used by Carslake, *et al* (2005), as being more desirable may still be able to detect differences in clustering.

2.5.6 Overall conclusions

Clustering of parasitic infections are variable and likely depend on a range of driving factors relating to the host, parasite and the environment. Here, I have analysed a range of parasites and given due consideration to their transmission mode and infection type (chronic versus acute). The key difference found in this study was between WMHV and its host, as well as several other parasites, primarily those transmitted via the environment. The close contact nature of WMHV and cowpox, as investigated by Carslake, *et al.* (2005) is comparable, and it is worth highlighting that both of these parasites show spatial clustering that is greater than the host. This similarity may provide some insight into the key aim of this chapter – to determine if transmission mode had implications for spatial clustering of parasitic infections. Furthermore, it appears that the spatial clustering of soil-transmitted parasites and flea-borne parasites (at least the *Bartonella* spp. analysed here) map onto that of the host relatively closely. As discussed above, this is not unexpected for the soil transmitted parasites presented in this study and, for vector borne parasites, given that clustering is consistent with the host, this may suggest that the fleas are not regularly undertaking large dispersal events.

Overall then, these analyses provide tentative support for the hypotheses that parasite transmission biology may leave a signal on the observed degree of parasite clustering relative to that of the host at different spatial scales, but in the absence of data on temporal (and therefore spatiotemporal) clustering it is hard to draw definitive conclusions. As such, any differences between the parasites in the wood mouse study system arising from any temporal differences in their viability away from the host remain unresolved.

3 SPATIAL SCALING OF WITHIN-HOST COINFECTION INTERACTIONS

3.1 Abstract

Interactions among coinfecting parasites are typically examined at the within-host level, often revealing strong effects on individual host susceptibility or disease progression. However the effect of these interactions on parasite transmission between hosts, and the spatial scales over which those effects operate, remain unknown. I analyse an extensive, spatially explicit dataset of the diverse community of parasites infecting wild wood mice in the UK, to assess the effects of local neighbourhood prevalence of each parasite species on individual-level infection risk by the other species, over an increasing range of spatial scales. My analysis revealed that the effects of within-host interactions between coinfecting parasites can indeed ripple out beyond the individual host, resulting in a network of facilitatory and suppressive effects on transmission among these parasites. However these between-host effects were only seen over relatively restricted distances around each host, over spatial scales likely reflecting the spatial scale of transmission. One implication of these effects may be the occurrence of knock-on, between-host consequences of antiparasite treatment for infection risk by non-target parasites, even for non-treated individuals.

3.2 Introduction

Hosts are typically infected by multiple parasite species throughout their lives (Cox, 2001). Interactions between coinfecting parasites have been identified in many human, livestock and wildlife systems (Nacher *et al.*, 2000; Graham *et al.*, 2001; Lello *et al.*, 2004; Ezenwa *et al.*, 2010; Telfer *et al.*, 2010; Babu and Nutman, 2016), with potentially important implications for host susceptibility, clinical disease progression, and treatment efficacy (Pedersen and Fenton, 2007). However, attention has focused primarily on the within-host mechanisms driving these interactions, with little understanding of the consequences they have for parasite transmission between hosts, and over what spatial scale this transmission interference occurs. Theory predicts coinfection interactions can alter transmission dynamics by affecting the population-level force of infection of either parasite species, but the nature of this scaling relationship can be highly non-linear, dependent on the mechanism driving the within-host parasite interaction (Fenton, 2008, 2013; Yakob *et al.*, 2013). Recent evidence from natural systems supports these broad predictions. For example, population-level analyses of nematode and protozoan infections in wood mice found no signal of any interspecific association (Fenton *et al.*, 2014), even though drug treatment experiments clearly showed these nematodes strongly suppress the protozoa within individual hosts (Knowles *et al.*, 2013; Pedersen and Antonovics, 2013, Chapter 2). Similarly, opposing individual- and population-level effects of helminth – *Mycobacterium bovis* (bTB) coinfection have been reported in African buffalo (Ezenwa and Jolles, 2015); anthelmintic-treated individuals benefited through reduced bTB-induced mortality, but this was predicted to increase bTB transmission at the population level due to the prolonged survival of, and hence opportunities for onward transmission from, bTB-infected hosts (Ezenwa and Jolles, 2015). These results suggest that within-host coinfection interactions can have potentially counterintuitive effects on parasite transmission through the host

population. However, there has been no explicit empirical demonstration of how within-host parasite infections affect parasite transmission in a natural system, nor quantification of the spatial scale over which such effects occur.

3.2.1 Examples of coinfection in natural systems

Coinfection is the normal state of a host in nature (Petney and Andrews, 1998; Cox, 2001). For example, the African buffalo (*Syncerus caffer*) hosts a number of parasite species, many of which have been shown to interact or co-associate within the host, such as strongyle nematodes and coccidial parasites, which have been shown to associate positively association (Gorsich, Ezenwa and Jolles, 2014). There are a number of other associations and interactions, positive and negative, in this system, including between *Mycobacterium bovis* (bTB) and strongyle nematodes (Jolles *et al.*, 2008), and between bTB and Brucellosis (Gorsich *et al.*, 2018a). These findings are important in highlighting the issue of the scale of the interaction; bTB is a risk factor for acquiring Brucellosis at the individual level, whereas at the population level, Brucellosis has a negative effect on bTB (Gorsich *et al.*, 2018b).

There are a number of mechanisms by which parasites can interact. For example, the strongyle nematode and bTB interaction appears to have a basis in the modulation of the host immune response (Ezenwa *et al.*, 2010). Similarly in humans, the nematode *Ascaris lumbricoides* provides a protective effect against cerebral malaria, via modulation of the IgE immune response (Nacher *et al.*, 2000). With different parasites stimulating differing immune responses (e.g. T-helper Type 1 versus T-helper type 2 responses), hosts may be forced to response in a way that may be beneficial to some parasites but costly to others. Another means by which parasites may interact with one another within the host is resource competition. For example, blood feeding worms have been

shown to compete with malaria parasites for red blood cells (Budischak *et al.*, 2018). Clearly, all these processes play out within individual co-infected mice, affecting their probability of being infected and/or the duration of infection. To what extent these within-host processes translate to affect the infectiousness of those individuals and the subsequent spread and infection risk among the wider host population currently remains an open question.

3.2.2 A natural model of coinfection: the wood mouse parasite system

The wild wood mouse (*Apodemus sylvaticus*) is native to the United Kingdom and much of Western Europe. These mice are host to over 30 species of parasites, from a range of taxa, and with a range of transmission modes, and hosts are rarely infected with only a single parasite species (Knowles *et al.*, 2013). As we know there is potential for ecological interactions between parasites (Pedersen and Fenton, 2007), and the diversity of parasites across this community makes these animals an ideal study system to investigate coinfection interactions between the different parasite species. In particular, experimental perturbation through anthelmintic treatment has demonstrated a strong, antagonistic, interaction between a nematode, *Heligmosomoides polygyrus* and a coccidian, *Eimeria hungaryensis* (Knowles *et al.*, 2013), such that anthelmintic-treated hosts had around a 15-fold increase in oocyst output by coinfecting *E. hungaryensis*. The previous drug-treatment experiment has clearly demonstrated that nematodes, and in particular the dominant species *H. polygyrus*, suppresses oocyst output of coinfecting *E. hungaryensis*. However the question remains whether the nematode-induced suppression of oocyst output, which is the vehicle for transmission of this species, affects *Eimeria* transmission between hosts, and over what spatial scale these effects may be observed.

3.2.3 Aims and hypotheses

To address the above question, and to explore the spatial scale of any between-host consequences of other potential within-host interactions between coinfecting parasites, I develop and undertake a novel analysis of spatially explicit data of wild rodents and their parasites, and show that within-host coinfection interactions do indeed influence between-host transmission, but only at local spatial scales.

3.3 Methods

3.3.1 Data

I analysed an extensive mark-recapture dataset of wild wood mice (*Apodemus sylvaticus*) and several of their parasites spanning multiple taxa and transmission modes, as discussed previously in Chapter 2 (Table 3.1; Chapter 2). This chapter uses data exclusively from grid HA1 in 2012. This grid was significantly larger than all others surveyed over the 6 year trapping period, and additionally, was subject to a much more intense trapping regime (every 2 weeks). This made it the most amenable option to investigate the spatial scale of coinfection interaction effects. There were 252 wood mice and a total of 986 independent captures. It is estimated that this represents approximately 80% of the total population of the grid (Fenton, 2020, Pers. Comm.). While there were multiple captures of individuals, each capture was treated independently given that each capture is temporally distinct, and the neighbourhood prevalence is determined by previous captures to a given temporal point relative to each spatial capture location (i.e., focal individuals at subsequent captures will have experienced additional exposure events than when captured previously, and so were treated independently).

Table 3.1. Parasites considered in the neighbourhood analysis study, with categorisation of potentially-important life-history characteristics (chronicity of infection and transmission mode). Overall prevalence of each parasite, across all grids and years, is also shown.

Parasite	Taxa	Infection Type	Transmission Mode	Prevalence
Heligmosomoides polygyrus	Nematoda	Chronic	Environmental	0.240
Eimeria hungaryensis	Protozoa	Acute	Environmental	0.147
E. apionodes	Protozoa	Acute	Environmental	0.157
Bartonella spp.	Bacteria	Acute	Vector (Flea)	0.478
Trypanosoma grosi	Protozoa	Acute	Vector (Flea)	0.118
Wood Mouse Herpes Virus (WMHV)	Virus	Chronic	Close Contact or Vector (Tick)	0.191

Using this dataset I first assessed how the infection risk of each parasite species associates with the neighbourhood prevalence of infection by that species around each focal individual, for increasing neighbourhood sizes. This helps inform how individual-level infection risk by each parasite species is related to levels of infection in the wider neighbourhood, and the spatial scale of those effects. Secondly, I assessed how the neighbourhood prevalence of one species affected the infection risk or intensity of other, potentially interacting, parasite species within neighbourhoods of increasing size. I did this first for two pairs of parasites where previous treatment experiments have quantified within-host interactions (or lack thereof) between them; the nematode *Heligmosomoides polygyrus* was found to suppresses the protozoan *Eimeria hungaryensis* (Knowles *et al.*, 2013), but had little impact on a different protozoan species, *E. apionodes* (Knowles *et al.*, 2013). Having validated the method with these known interacting/non-interacting pairs of species, I then extended the analysis to consider other pairs of species for which there is no current evidence of interaction between them.

3.3.2 Neighbourhood analysis

I considered each individual at a single capture point as a focal individual in turn. Around each focal animal at each capture, we defined its neighbourhood of size *r* as the trap locations within *r* metres of the focal animal's capture location (Fig. 3.1). I then determined the number and identity of every other individual caught at traps within that neighbourhood at any time point prior to the capture date of the focal individual; captures of animals within the neighbourhood but after the focal animal's capture date were ignored. I then calculated the neighbourhood prevalence of infection by a specified parasite as the proportion of neighbours that were positive for that parasite within the focal's neighbourhood of radius *r*.

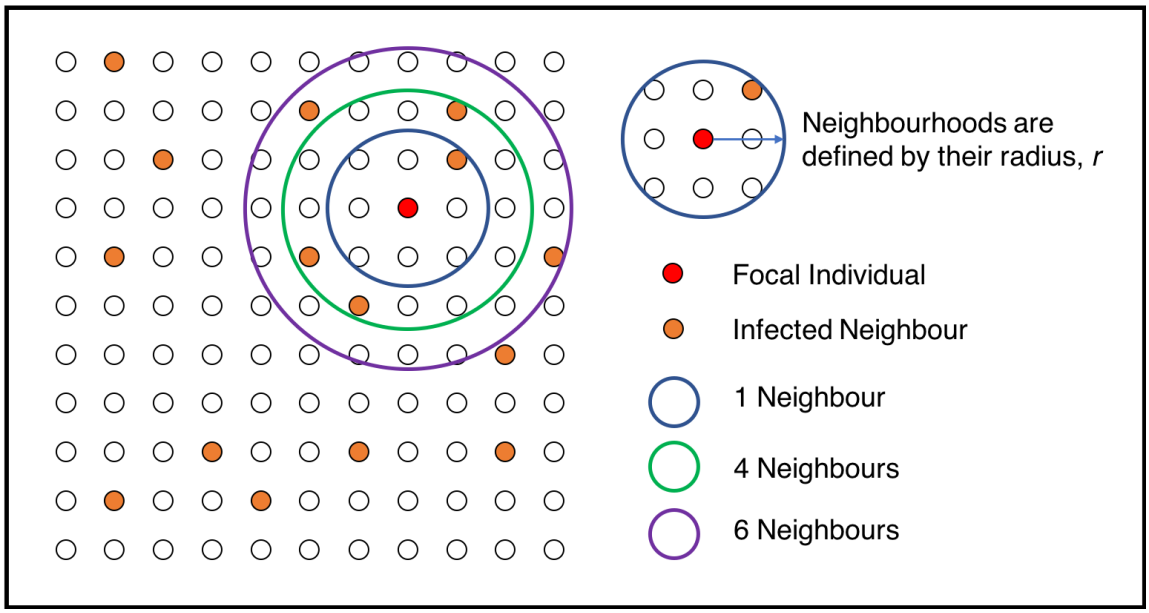

Figure 3.1. An example of 3 neighbourhoods constructed from a focal individual on a hypothetical grid. Each subsequent neighbourhood, larger than but encompassing the previous one, takes into account all neighbours in the neighbourhoods smaller to it, and any additional neighbours. Shown here are infected neighbours. The algorithm will also search separately for uninfected neighbours.

I compiled this dataset using all mice as focal animals in turn, for all possible neighbourhood sizes, from $r = 10$ m (immediate neighbours only) up to $r = 100$m (the entire grid). For some individuals and neighbourhood sizes, it was not possible to

account for their entire neighbourhood as some of it lay outside the experimental grid. In these instances, we assumed that the grid was representative of the wider woodland, and as such the observed prevalence of infection within the neighbourhood on the grid reflected that of the focal's entire neighbourhood.

3.3.3 Statistical Analysis

My overall aim was to establish whether within-host interactions between coinfecting parasites affect the force of infection by one of those parasite species on other mice within a specified neighbourhood size. I did this by examining whether the infection intensity (egg/oocyst burden of infection of a specified parasite among infected animals) or when this was unavailable, the infection risk (presence/absence of infection by a specified parasite) in each focal host varied with the prevalence of infection by other, specified parasites within neighbourhoods of each specified size. For each neighbourhood size (r = 10m to 100m), I ran Generalized Linear Mixed Effects Models (GLMMs) implemented in a Bayesian framework using stan (rpackages: rstan (version 2.17.3), brms (version 2.3.0)). I did this for all possible of pairs of focal and potentially-interacting parasites in the dataset. All analyses were undertaken in R (version 3.5.0).

Each individual's focal parasite status was measured either as intensity (Gaussian model, with egg or oocyst output per gram of faeces among infected animals only, for those parasites where this was possible [*E. hungaryensis, E. apionodes* and *H. polygyrus*]), or as infection presence/absence (binomial model with log link, for WMHV, *Bartonella spp.* and *T. grosi*), as my response variable. My main fixed effect of interest was parasite prevalence of each specified parasite within the defined neighbourhood. In all models I also controlled for the following potential confounders of the focal individual: its infection status (positive/negative) with the potential interacting parasite, sex (male or

female) and age (adult or not). I also controlled for the total number of animals in the defined neighbourhood, and the date of capture as a 2nd-order polynomial to account for non-linear seasonal effects, and focal ID number as a random effect to control for potential pseudo-replication arising from multiple captures of the same individual. I did not carry out any model reduction or simplification, to ensure consistency in model structure across all analyses, and I report the coefficients (median model estimate, ± 95% credible intervals) for neighbour parasite prevalence on focal host infection intensity or presence/absence for each neighbourhood size (separate analyses run for *r* = 10m to 100m), while controlling for the same set of potential confounders in all analyses.

3.3.4 Accounting for treated animals

The dataset contained some animals that had been treated with the anthelmintic Ivermectin or the anticoccidial drug Vecoxan. Treated individuals were excluded as focal individuals in this study as it may confound the results. However, they were allowed to be considered as neighbours (i.e., they do occur on the grid, and so could potentially contribute to infection of focal individuals). Ivermectin has been shown to be effective in reducing the intensity of nematode infection in wild wood mice, but the effect is very transient, with treated individuals reinfected within days of treatment (Knowles *et al.*, 2013; Clerc *et al.*, 2019). Vecoxan (an anti-Coccidial drug) was administered to some animals on the grid, however there is no observable effect of Vecoxan on either *Eimeria* species (M Clerc, 2017). Given the lack of observable effect, I ignore any Vecoxan treatment, and consider Vecoxan treated animals as any other focal host in our analyses.

3.4 Results

3.4.1 Single parasite models

Initial analysis examined how neighbourhood prevalence by each parasite associated with focal infection risk by that same parasite species, at increasing spatial scales. These are displayed in Fig 3.2 along the diagonal of the panel. Infection risk by *E. apionodes* and *H. polygyrus* were each positively associated with neighbourhood prevalence of those parasites, up to spatial scales beyond 2 home ranges of the host (~30-50m). WMHV prevalence appears to become more positively associated with increased risk of infection at the largest spatial scales, close to the size of the grid as a whole. There appears to be no notable association between *T.* and *Bartonella spp.* prevalence and infection status of focal individuals across any spatial scales. Interestingly, *E. hungaryensis* appears to be negatively affected by prevalence at small spatial scales.

3.4.2 Coinfection interaction models

These results are presented in Fig 3.2 in the off diagonal plots. My analysis revealed a general reduction in individual-level intensity of *E. hungaryensis* in focal hosts with increasing *H. polygyrus* prevalence in the wider neighbourhood, but primarily at spatial scales which approximate the mean home range size of wood mice (~14m radius) (Carslake *et al.*, 2005). At spatial scales larger than 2 home ranges, there was no association between neighbourhood *H. polygyrus* prevalence and *E. hungaryensis* intensity. Conversely, I found no effect of *H. polygyrus* neighbourhood prevalence on *E. apionodes* intensity, the species known not to interact strongly with *H. polygyrus*) at any spatial scale. Hence increasing neighbourhood prevalence of *H. polygyrus* appears to result in a reduction of *E. hungaryensis* infection risk, the species which *H. polygyrus* is known to suppress oocyst output of within individual hosts, but this is only seen in

neighbourhoods up to around 1 host home range diameter. Furthermore, there was no detected effect of neighbourhood *H. polygyrus* prevalence on *E. apionodes* infection risk, the species known not to interact strongly with *H. polygyrus* within individual hosts. Hence, these results provide proof of concept that this technique is a viable means of detecting parasite-parasite interactions (and non-interactions) across a range of spatial scales using observational data.

Having shown that parasites with strong within-host interactions leave a signal of that interaction at localised spatial scales, and that parasites that don't interact do not leave such a signal, I sought to test for evidence of interactions within the wider community of parasite species infecting these hosts, using the directly-transmitted wood mouse herpesvirus (WMHV), the flea-borne bacterium *Bartonella spp.,* and the flea-borne protozoan *Trypanosoma grosi.*

I found negative associations between neighbourhood *H. polygyrus* prevalence and individual-level infection risk of both WMHV and *T. grosi* for neighbourhood sizes up to around 36m. The opposite was seen for the association between *H. polygyrus* prevalence and *Bartonella spp.* infection risk over similar neighbourhood sizes, with *Bartonella spp.* infection risk increasing with neighbourhood *H. polygyrus* prevalence for neighbourhoods up to ~40m. There was also a positive effect of *E. apionodes* prevalence on *T. grosi* infection risk across neighbourhood sizes up to ~50m, and reciprocal positive effects of neighbourhood *Bartonella spp.* prevalence on WMHV infection risk, and of WMHV neighbourhood prevalence on *Bartonella spp.* infection risk across neighbourhoods of 50m and over. There were no obvious associations between neighbourhood prevalence of either *E. hungaryensis* or *T. grosi* on infection risk of intensity of any other parasite.

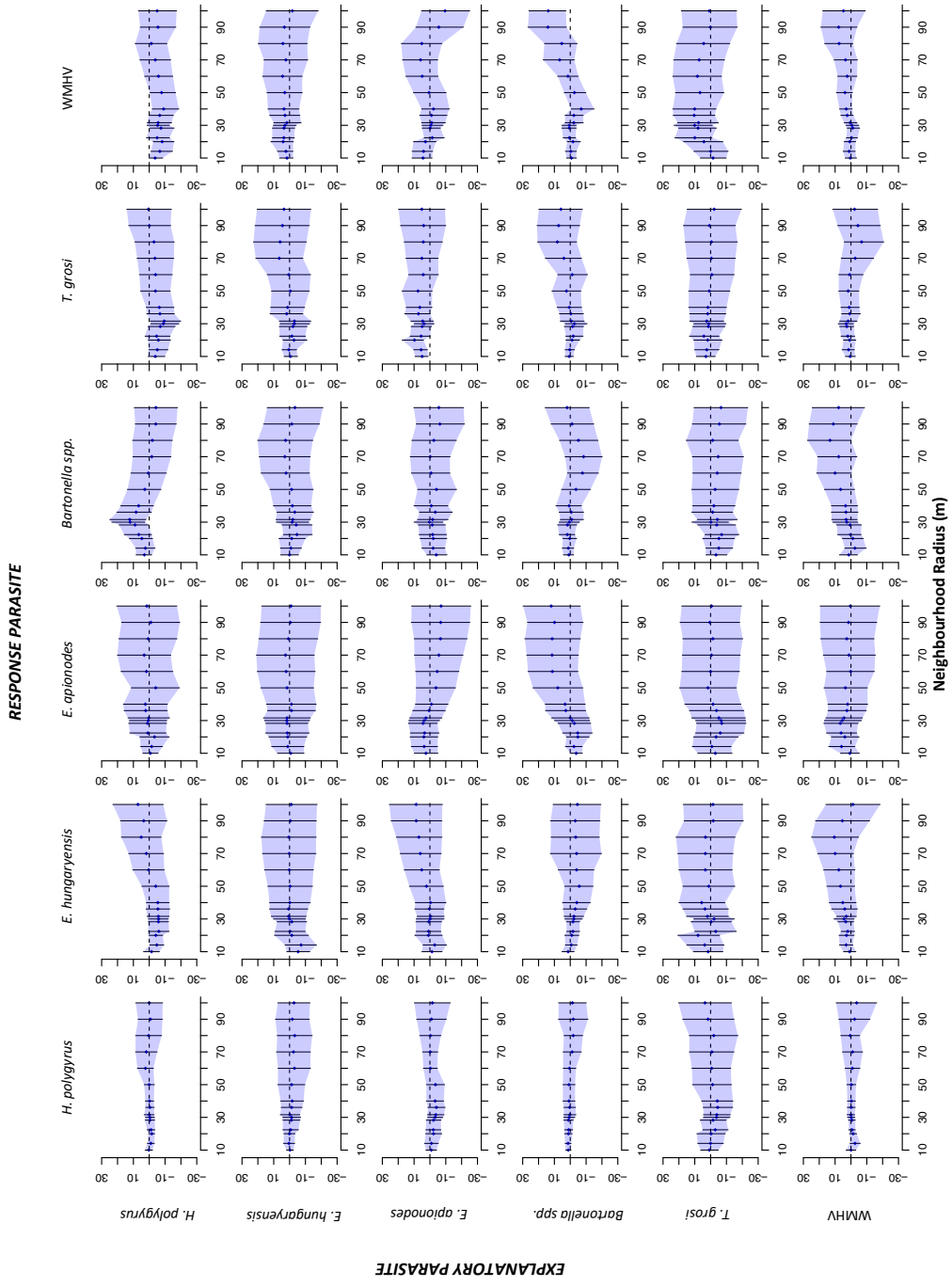

Figure 3.2 Plots of neighbourhood effects of the explanatory parasite, listed on the left, on the response parasite, listed on the top, for increasing neighbourhood sizes (radius in metres around focal individuals) on the x-axes. The diagonals show the effect of each parasite on itself. Each plot shows outputs from 15 separate models for a total of 540 models in this figure. Points show median model estimates and envelopes show 95% credible intervals. Interactions of note are indicated with a *.

3.4.3 Null Model Simulation

To help assess the ability of this approach to detect true relationships and to ensure that the method is not biased to generating associations when there are none, simulations were done with an absence of coinfection interactions. Based on the prevalences of each parasite in the *H. polygyrus* – *E. hungaryensis* relationship, infections were randomly assigned to animals in the dataset, while maintaining the spatial coordinates of the real animals. Upon undergoing further neighbourhood analysis to ensure that the relationship does not emerge via some statistical artefact, there was no neighbourhood size which this was the case (Fig 3.3).

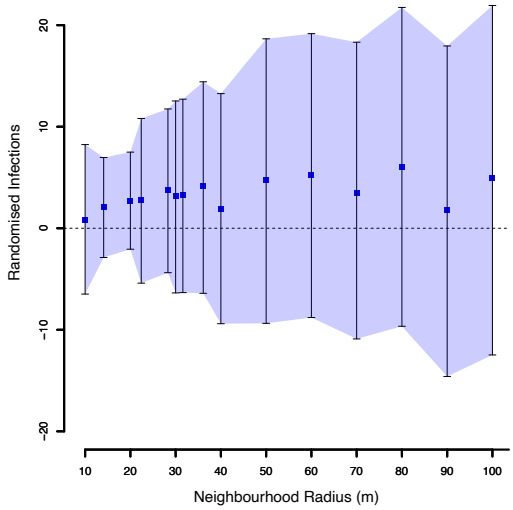

Figure 3.3 Neighbourhood effects of a randomised parasite (based on *H. polygyrus* prevalence) on *E. hungaryensis*. The diagonals show the effect of each parasite on itself. Each plot shows outputs from 15 separate models for a total of 540 models in this figure.

Sample size here represents approximately 80% of the population (Fenton, 2020, Pers. Comm.). Given the ability of the technique to detect the known *H. polygyrus* – *E. hungaryensis*, as well as not detecting an interaction in randomly simulated data, this sampling effort would seem sufficient.

3.5 Discussion

3.5.1 Between host consequences of coinfection

The within-host dynamics of coinfection, and how these alter disease progression, have been well documented in a range of systems. Here I have been able to demonstrate, using one of these well-characterised interactions (Knowles *et al.*, 2013), that there are detectable consequences of within-host interactions for between-host transmission. Given this suppressive interaction of *H. polygyrus* on *E. hungaryensis*, and the 15-fold reduction in *E. hungaryensis* oocysts shed by *H. polygyrus* coinfected hosts (Knowles *et al.*, 2013), it would seem only logical that the transmission potential of *E. hungaryensis* is severely decreased. I show that the effect of this negative interaction is detectable up to two home ranges around the host, with strongest effects being seen at distances closest to a single home range (Fig 3.2). Not only does this support the expectation that there would be consequences for transmission, the scale of transmission interference detected is sensible given the biology of transmission of these parasites. Consider two hosts, individual *a* and individual *b*, each with a home range around them. There can be a degree of overlap of this home range, such that *a* and *b* occupy some overlapping space. This is important, as given the environmental transmission via infectious propagules in the soil for *H. polygyrus* and *E. hungaryensis*, there must be some degree of spatial overlap of habitat use by the hosts for transmission to occur. At the most extreme case of a lack of spatial overlap, with only the smallest shared space usage, the maximum distance from the far edge of *a*'s home range to the far edge of *b*'s home range is approximately 2 home ranges in size – aligning with the scale of interaction that I have seen in the results of this study.

This validation of the method has allowed a much more thorough exploration of other potential parasite interactions. I have detected 6 novel interactions at scales beyond the host. Strikingly, I found *Bartonella spp.* (Fig 3.2) positively and *T. grosi* (Fig 3.2) negatively affected by *H. polygyrus* neighbourhood prevalence. While the direction of these effects are opposing, the peaks of where the spatial scale was most strong is the same, between 1 and 2 home ranges. This is particularly interesting as both parasites are flea transmitted, suggesting that there may be a link between transmission mode and the spatial scale over which any interactive effects occur.

3.5.2 Methodological advance in the detection of coinfection interactions

This technique represents a significant move forward in the detection of coinfection interactions from observational data. Crucially, I have shown that the spatial scale of observation is vital in detecting these interactions, given the seemingly local nature of many that I have detected. This sensitivity of detection to the spatial scale examined may explain why many larger-scale, population-based analyses of coinfection interactions fail to discover any significant association between parasites from cross-sectional data (Fenton *et al.*, 2014). The results presented here show that examining whole-population data effectively averages across all spatial scales, thereby obscuring the local processes of transmission interference arising from within-host interactions that play out over much smaller spatial scales around each individual host.

It should also be noted that the form of rodent mark-recapture experiment on regularly spaced grids used in this analysis is not unique to these data; many studies of wildlife use similarly collected data (Turner *et al.*, 2014). Given the utility of the technique, and the pre-existing data available, application of neighbourhood analysis to both new and

legacy data of this form may provide significantly more insight into the interactions between parasites, and therefore their community ecology.

While this technique has been designed as a bespoke analysis tool for this data set, conceptually it need not be restricted to mark-recapture data on a regularly spaced grid. The concepts of neighbourhood analysis as presented, that parasite-parasite interactions can be detected at intermediate scales by considering neighbours of a host across a range of scales could be applied to other systems and types of data. This could lend itself to investigations of parasite-parasite interactions across multiple scales in a variety of systems.

The temporal dimension of this technique has not been fully explored in this chapter. However, this is ripe for further exploring the spatiotemporal interactions between parasites. The neighbourhood analysis algorithm allows for selectively altering the temporal window in which neighbourhood prevalence is calculated. This opens up possibilities for exploring other related questions. For example, investigating whether there is a defined temporal window in which the neighbourhood prevalence is important, such as how this relates to environmental viability of the parasite.

Of the self-associations investigated in this chapter, the most striking is that of *E. hungaryensis*. One would not expect a negative effect of neighbourhood prevalence on focal infection risk at any spatial scale; positive associations should be much more likely, due to localised transmission processes, resulting in clusters of infection. However, given the negative trend between neighbourhood *E. hungaryensis* prevalence and individual-level infection risk across a number of models at small spatial scale (Fig 3.2), this affect seems consistent. One possibility is that there is some form of long-lasting

protective immunity to *E. hungaryensis* in the wild, and that a high local prevalence is indicative that an individual may have been infected at some point in the recent past and is resistant to subsequent infection. How likely, or effective immunity to *E. hungaryensis* is in the wild is unclear, however (Clerc, 2017).

3.5.3 Overall conclusions

Overall I have shown that within-host interactions between coinfecting parasites can affect their localised transmission dynamics, leaving a signal of that interaction at spatial scales beyond the individual host. Furthermore the spatial extent of that effect operates across different spatial scales for different parasites, and is likely to reflect their spatial scale of transmission. There are many human, livestock and wildlife disease systems where there are well established within-host parasite interactions among coinfecting parasites, and a growing body of research has examined the consequences of those interactions for the success and impact (beneficial or detrimental) of disease treatment approaches on individual host health (Griffiths *et al.*, 2011, 2015). However, my results suggest that there could be knock-on, between-host consequences of such treatments, particularly in communities experiencing high coverage mass drug administration, for localised transmission dynamics of non-target parasites, with implications for infection risk even among non-treated individuals.

4 SCALING OF COINFECTION INTERACTIONS ACROSS SOCIAL NETWORKS

4.1 Abstract

Classical models of infectious diseases generally assume random mixing of individuals in a population, like a chemical reaction. In these models each individual is as likely to encounter every other individual equally. While in some cases, these models adequately describe the dynamics of the infection across the host population, ignoring heterogeneities in contacts between individuals can overlook a significant element of the transmission biology of the pathogen, driven by variation in host behaviour. In the last decade or so, studies focusing on how social networks relate to the spread of infection have increased dramatically in number. Using a well-studied population of African Buffalo from Kruger National Park in South Africa, I explore how two measures of an individual's place in a social network, eigenvector centrality and degree, affect the disease status of individuals. I find that for some parasites, eigenvector centrality affects disease status but that for all parasites degree has no effect. I then adapt the neighbourhood analysis technique presented in chapter 4 to investigate potential novel parasite-parasite interactions. I have detected a previously unknown association, with coccidia positively predicting bTB infection status. This adaptation of neighbourhood analysis has shown that is an additional tool in the search for ecological interactions between parasites.

4.2 Introduction

4.2.1 Assumptions of homogenous mixing and its limitations

Typically, models of the transmission and spread of infectious diseases in their host populations make the assumption of homogenous mixing, whereby contacts between susceptible and infectious individuals occur at random, and all individuals in the population are equally likely to encounter each other (Begon *et al.*, 2002). This assumption of homogenous mixing has proven effective in predicting the spread of infectious diseases within some systems. That being said, these models fail to capture potentially significant elements of host biology that can influence pathogen transmission and spread. Those models typically partition the host population into discrete, homogenous categories, within which all individuals are treated as equal. In the case of the simplest *Susceptible-Infected-Recovered* (SIR) models, individuals are assigned to a class based on their infection status, in this case *S* for susceptible (uninfected), *I* for infected and *R* for recovered. The changes through time in the numbers of individuals within each class are then typically described by a system of ordinary differential equations which can be used to model disease progression numerically (in all but the most simple of models), or to calculate metrics such as R_0 (the basic reproductive number) analytically. A key component of these models is the transmission rate, usually represented by β, which is a composite parameter, describing the rate at which susceptible and infected individuals randomly encounter each other, and the probability of pathogen transmission given an encounter (Begon *et al.*, 2002; McCallum *et al.*, 2017).

While the assumption of homogenous contact structures makes the corresponding models tractable and accessible, there is a significant amount of heterogeneity in host biology that is not realised in these models. For example, individuals may differ behaviourally, such that some interact more with other individuals, and as such be potentially at greater risk of being infected or infecting others. The spatial arrangement of hosts may also be heterogeneous, with some clusters and sub-communities giving rise to some individuals having a higher frequency of contacts with each other than with others. There are a number of adaptations to the compartmental framework which have aimed to account for heterogeneities and non-linearities in transmission. For example, frequency dependent transmission assumes the rate of acquisition of new infections is dependent on the frequency or proportion of infectious cases in the population, rather than on absolute population density (Begon *et al.*, 2002). Hence, per capita contact rates are assumed to be constant, independent of density, and has been traditionally used to model vector borne or sexually transmitted infections. One would not expect the number of sexual partners to increase linearly with density, as individuals find or have a fixed number of sexual partners regardless of the population size. Neither would one expect the number of vector bites to increase linearly with host population size, as vectors require a limited number of blood meals and, for example, flying vectors (e.g., mosquitoes or tsetse flies) may be sufficiently mobile to locate the appropriate number of hosts locally, regardless of overall population size. Thrall, Antonovics and Hall (1993), demonstrate that frequency dependent models can predict coexistence between host and parasite in the case of sexual or vector transmission. Nonlinearity may also be present in the transmission function itself. These nonlinearities represent the underlying biology of the host and parasites, while still being captured in the compartmental framework (Fenton *et al.*, 2002). Indeed, this apparent *flipping* from density to frequency

dependence has been modelled for the virus of wild rodents, cowpox (Smith *et al.*, 2009). We can therefore think of frequency and density dependence being the two extremes of a transmission function that varies with density in this instance. As density increases, from a low density starting point, we see what appears to be density dependence, and then when densities increase while the population is at a higher density, we see something more akin to frequency dependence (Smith *et al.*, 2009).

4.2.2 Social network approaches and network features

An alternative to the above compartmental models is to treat individuals as individual components and describe contacts between them explicitly; in other words, to consider the social network of contacts between individuals across the population. Social networks explicitly represent potential pathways for contact between individuals, and hence routes by which disease transmission can occur assuming transmission is by direct host-to-host contact, or through a shared environmental medium (White, Forester and Craft, 2017). For example, using genetic data from *Escherichia coli* in giraffes, VanderWaal *et al.*, (2014), demonstrated that the transmission network (i.e., the links through which the pathogen passed from individual to individual) closely correlated to the social network (i.e., the observed occurrences of individual-to-individual contacts). The dependence of transmission on social structure highlights the potential importance of variations in social network structure, and how they may serve as powerful predictors of infection.

The aim of developing social network representations of populations is not to discount the value and utility of compartmental models, but rather to understand fully the importance of heterogeneities in contact structure. Social networks have been used to investigate the connectedness and potential for pathogen transmission in populations

(Hirsch *et al.*, 2013) and to successfully predict the likelihood and intensity of infection (Godfrey *et al.*, 2010). They have also been used to characterise how variation in host contact types can have implications for likelihood of disease transmission (Blyton *et al.*, 2014). While *Escherichia coli* is primarily spread through the faecal-oral route, Blyton *et al.*, (2014) found that spatial proximity was less explanatory than host contacts, showing a difference in what types of contact a host has, and its risk of infection. These approaches have been employed in empirical and modelling studies of infectious disease transmission across a wide variety of systems, from reptiles (Godfrey *et al.*, 2009; Aiello *et al.*, 2014), ungulates (VanderWaal *et al.*, 2014), marsupials (Corner, Pfeiffer and Morris, 2003), primates (Griffin and Nunn, 2012; Carne *et al.*, 2014; Romano *et al.*, 2016) and humans (Bansal, Grenfell and Meyers, 2007; Salathé and Jones, 2015).

Regardless of the system being described, all social networks take the same basic structure. Individuals are represented in the network by *nodes* (White, Forester and Craft, 2017). A connection between two nodes is called an *edge* (White, Forester and Craft, 2017). Edges can be directional (i.e. information is flowing in one direction) or non-directional (i.e. information flows evenly between nodes). These edges can be *weighted* or *unweighted*, meaning that some edges may represent more information flow, or occur more often, than others (White, Forester and Craft, 2017). The structure and composition of the network can then be measured in various ways. The *distance* is the number of edges between any two nodes (White, Forester and Craft, 2017). How connected an individual node is can be measured by a range of measures of known as *centrality*. *Degree centrality* is the number of edges leading to or from a node (White, Forester and Craft, 2017). More comprehensive than degree centrality is *eigenvector centrality* which is a much more holistic measure than degree, given that it not only accounts for the number of edges attached to a node, but also the relative importance

of this node in the network as a whole (White, Forester and Craft, 2017). The decision of which metric(s) to use is dictated by the amount of data available to inform them and the questions being addressed.

4.2.3 Network neighbourhood analysis and the spatial scale of coinfection interactions

It may be easy to consider the utility of network analysis in species with limited contacts, such as the territorial mice described elsewhere in this thesis. However, what of those species that are drastically different, and significantly more social?

As explored in chapter 3, I have developed a method for looking at how the neighbourhood context of infection and coinfection influences an individual's infection risk, over intermediate spatial scales between the individual, and the whole population, termed neighbourhood analysis. This was developed for, and successfully tested on, the spatially explicit dataset of the territorial wood mice. I am interested in testing whether this method can be extended to ask questions of transmission and coinfection interaction in a very different system spatially and socially, and determine whether there is spatial or social scaling of these processes. Herding species present a very different social situation compared to territorial species such as wood mice, due to their close association with each other. For such animals, by observation of known individuals within herds, association data can be generated and subsequently a social network of potentially transmission-relevant contacts can be constructed.

Group living is often associated with the cost of higher parasitism, due to close contacts between individuals and high (sub-)population densities (Altizer *et al.*, 2003). However, not all individuals in groups necessarily behave the same, nor will they interact homogenously within those groups. Some individuals will inevitably contact more with other individuals within and between groups, thereby acting as potential super-

spreaders, driving transmission dynamics among the group. Alternatively, other individuals may have very limited connections with other members of the group, thereby acting as potential sinks of transmission, with limited opportunities for onward transmission. How the group is structured may therefore provide an insight into how the parasite dynamics are shaped.

In this chapter I analyse association data within the African buffaloes system. These animals have several micro- and macroparasites, and there are a number of well characterised interspecific interactions occurring between these parasites. Notably, there is an interaction between bovine tuberculosis (*Mycobacterium bovis,* bTB) and strongyle nematodes, with nematodes and bTB being negatively associated at both the within-herd and whole population levels (Jolles *et al.*, 2008). However, at the individual level, when there is an absence of nematode infection, for example through anthelmintic treatment, bTB fails to infect the host, and this has been attributed to immunomodulation by the nematodes. Nematodes stimulate the T-Helper Type 2 (Th2) component of the immune system, which is antagonistic to the T-Helper Type 1 (Th1) response that combats some microparasites such as bTB. Hence, when nematodes are removed by deworming, the Th1 response increases, preventing bTB infection (Ezenwa *et al.*, 2010).

In addition to the bTB-nematode interaction, there has also found to be a positive association between nematodes and coccidia. This is irrespective of age, sex or season (Gorsich, Ezenwa and Jolles, 2014). Common exposure to both parasites was considered one means of explaining this correlation. However, there are well characterised negative interactions between nematodes and coccidia in other systems (e.g. Knowles *et al.*, 2013; Pedersen and Antonovics, 2013) and there may be a number

of underlying reasons for this correlation. Furthermore, several microparasites of these buffaloes have also been found to interact with each other. bTB is a risk factor for the acquisition of brucellosis at the individual level, however the reverse is not true, with there being no effect of brucellosis on bTB infection (Gorsich *et al.*, 2018). At the population level though, brucellosis has a negative effect on bTB, with there being no effect of bTB on brucellosis at this scale (Gorsich *et al.*, 2018a). Population level competition, as seen here, is likely a consequence of the immunosuppressive effects of bTB operating at the individual level.

It is clear from these examples, and my work in chapter 4, that the scale of observation is important in determining the observed effects of a parasite-parasite interaction (Jolles *et al.*, 2008; Gorsich *et al.*, 2018b). These studies on the buffalo have given insight into the effects of parasite-parasite interactions across a range of ecological scales (individual, herd and whole population), however, more can be done to understand the intermediate scales (i.e. beyond the individual but within the herd/population), and in particular the spatial scales over which the effect of those interactions can be observed. As stated above, Chapter 3 introduced a novel method termed neighbourhood analysis to analyse infection and coinfection effects at intermediate spatial scales. The aim here is to adapt this method to apply to social networks, to examine the scale of infection and coinfection interaction effects in the buffalo system. While wood mice are territorial, buffalo are significantly more social animals. As such, I can use a social network approach as the basis for this neighbourhood analysis and use distance between individuals on the network as a proxy for overlap of space use by individuals. Therefore, within herds, I can span the individual and whole herd scales in an incremental manner to understand how infection and coinfection interactions scale spatially across the social network.

4.2.4 Aims and hypotheses

In this chapter I investigate the questions raised in this chapter using the African buffalo social network relate to individual levels of parasitism. For example, are more connected individuals, by either degree or eigenvector centrality, more likely to be infected? Is this true for all parasite species? Secondly, I determine how infection risk scales with distance in the network by applying a form of neighbourhood analysis to these data. Finally, this will be used to investigate the spatial scaling of known parasite-parasite interactions, and to search for new potential interactions.

4.3 Methods

4.3.1 Data

Dr Vanessa Ezenwa of the University of Georgia has kindly allowed access to the following dataset to undertake the analyses for this thesis. Female African Buffalo were trapped between June 2008 and August 2012. Trapping was conducted in the southern region of Kruger National Park in South Africa. 200 animals were captured in 2008 by helicopter which were then fitted with radio collars. They were trapped twice yearly after this. They were assigned to treatment or control groups randomly upon their initial capture. Lost individuals, as a result of emigration or death were replaced, with their treatment status matching that of the individual they replaced. Animals were from herds in two locations, Lower Sabie, and Crocodile Bridge. Their herd was recorded at subsequent captures. Initial estimates of the Lower Sabie herd population was 1117. The Crocodile Bridge herd was estimated to be 2104. 100 animals in each herd (50

control, 50 treated with slow-release fenbendazole bolus) were enrolled in this study. Animal age was recorded and faecal and blood samples were collected to test for micro- and macro-parasites. For more information on the African buffalo dataset, please see (Ezenwa and Jolles, 2015).

4.3.2 Social Network Analysis

Social network analysis was conducted in R (version 3.5, R Core Team, using the package igraph (version 1.2.1). The data were subsetted for year, and season (dry or wet) within each year. Interaction matrices were generated using a custom built function, which took each possible pair of individuals in turn, and determined if at any time within that year and season combination, they were recorded as occurring within the same group. If they were, the pair were assigned a 1, or if they were not, they were assigned a 0. The resulting dataset across all pairs was then subsetted to exclude all pairs with a 0, to produce a matrix of all edges in the network. igraph was then used to convert these interaction matrices to graph objects which could then be used to extract a number of individual network characteristics.

4.3.3 Statistical models of individual network metrics

The first set of models assessed how individual infection status was influenced by individual network metrics (degree and eigenvector centrality), which provide information on how well connected each individual is within the larger network. This analysis used a series of Generalized Linear Mixed Effect Models (GLMMs) conducted using stan, in the R package brms. These models sought to determine the effect of those individual level network features on whether an individual was infected with a

given parasite or not, using a Bernoulli error structure. Given that some animals in the dataset had received anthelmintic treatment as part of an on-going study, which may affect observed infection levels, only the infection status of individuals that were untreated were considered in these models. However, all animals were considered when generating the social networks that were used to determine the individual network metrics.

The structure of these statistical models was as follows:

> Parasite Infection Status ~ Eigenvector Centrality + Degree + Age + Year + Season + Herd + Treatment + (1| Animal ID)

All parasite infection status is recorded as 0/1, presence/absence data and so the models will have a binomial error structure. Eigenvector centrality and degree were standardised (variable – mean variable / standard definition variable) for inclusion in the model. Each model included each individual's eigenvector centrality and degree as the main variables of interest in predicting infection risk for each individual, and the animal's Age (continuous variable, in months), and the Year (factor: 4 levels, 2009-2012) and Season of sampling (factor, 2 levels, Wet or Dry) were included to account for any variation that may occur as a result of these. In addition, all models included each animal's unique identification number as a random effect to control for pseudoreplication arising from multiple samples from each individual. In all analyses, there was considered to be strong support for a given variable if the 95% credible intervals of its estimate did not cross zero.

There is, of course, some degree of correlation between degree, and eigenvector centrality across nodes in the network (~69%). This make sense, since nodes with a high number of edges are more likely to be important, but that ultimately, low degree

individuals attached to these high degree individuals will have a higher eigenvector centrality as a consequence of this connecting edge. Conversely, high degree individuals connected to a number of unimportant nodes will have lower eigenvector centralities as a consequence. While there is correlation, this was considered within the bounds by which both terms could be included in the same model (< 70%), making it possible to directly compare the relative influence of these two related, but different metrics (Harrison *et al.*, 2018). That being said, I have also elected to run statistical models looking at either degree or eigenvector centrality independently to ensure that any results are robust.

4.3.4 Neighbourhood Analysis of Social Networks

The approach described in chapter 4 to quantify the influence of neighbourhood context on individual infection risk from spatially explicit data was adapted to be used on social networks. Instead of Euclidian spatial distance, the distance from each focal individual in the network is used. To explain the process, consider a focal individual. This individual will have "neighbours", or nodes associated with it via an edge in the network. Hence individuals directly connected to that focal individual by one edge can be thought of as its immediate neighbours (within 1 edge 'distance'), whereas those individuals connected to the focal via an immediate neighbour can be thought of as a neighbour within 2 edges 'distance' etc. For neighbourhoods of 1 edge in distance, I determined the infection state of those neighbours, and then calculated the prevalence (proportion of neighbours infected with the parasite of interest) in this neighbourhood of the focal individual. Then, for neighbourhoods of 2 edges in distance (i.e., not only the nodes connected to the focal, but also the nodes connected to those nodes that are 1 edge from the focal), were considered when determining the prevalence. This was done incrementally for all distances *D*, from 1 up to the maximum distance from the

focal. The largest network distance in the dataset was 7, although for some season and year combinations it was as small as 3. Neighbourhood prevalences were calculated both for the focal parasite in a given analysis, to determine the effect of neighbourhood prevalence on its own infection status, as well as for all possible interacting parasites in the neighbourhood. Once these neighbourhood prevalences had been calculated, GLMMs were conducted, using the format:

> Focal Parasite ~ Prevalence of focal parasite at distance *D* + Age + Season + (1 | Animal ID)

> Focal Parasite ~ Prevalence of Potential Interacting Parasite at distance *D* + Focal Status with Potential Interacting Parasite + Age + Season + (1 | Animal ID)

Where *D* is the distance up to which we consider the neighbourhood.

4.4 Results

4.4.1 Network structure over year and season

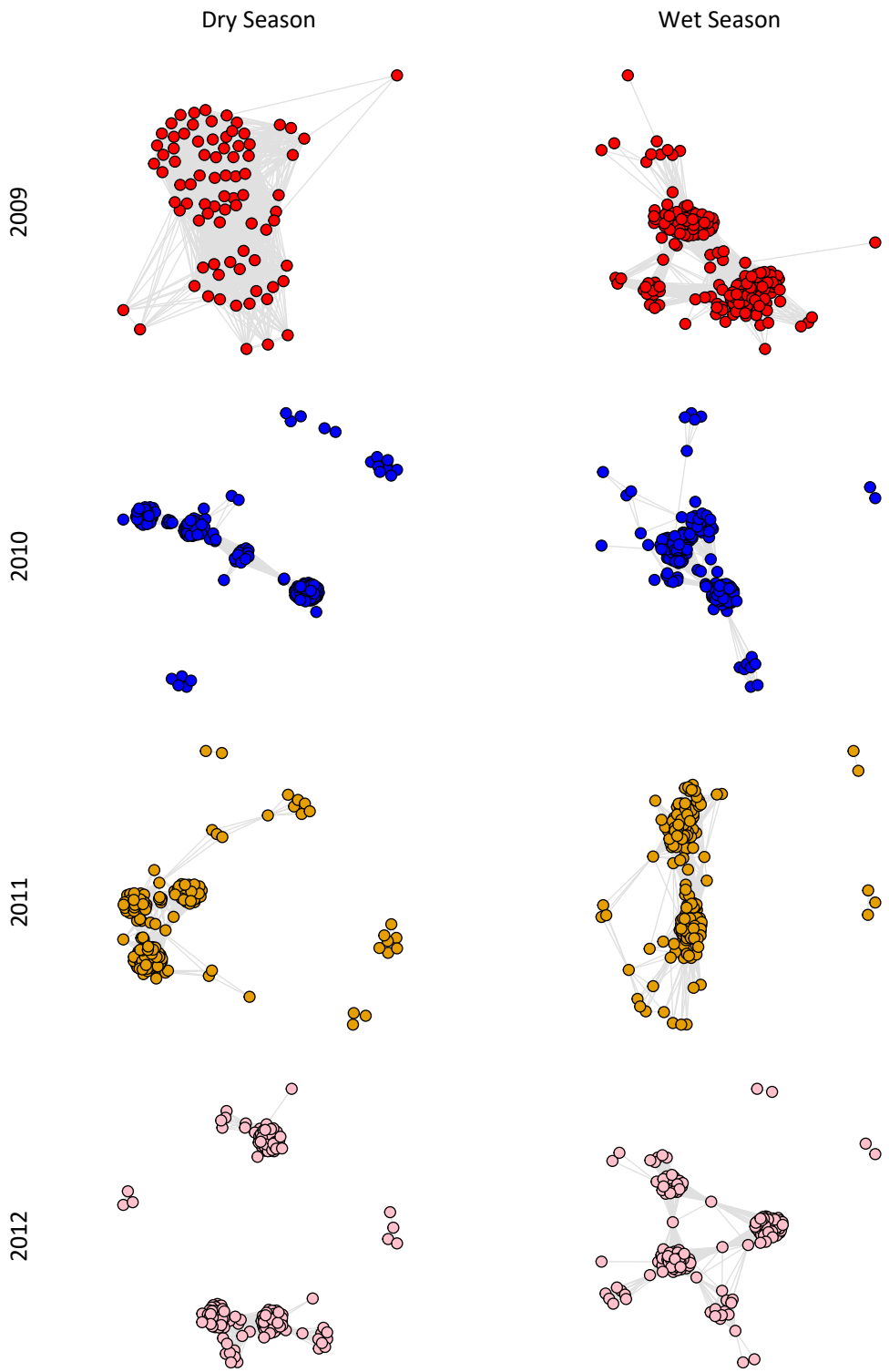

Figure 4.1 Network plots for each year and season combination.

4.4.2 Effects of individual network metrics on individual infection risk

Standardised degree and standardised eigenvector centrality histograms are shown in fig 4.3. Degree was not a strong predictor for the presence of any of the parasite species (the 95% credible intervals crossed 0 for the effect of degree on all parasites), however for models including both variables, there was a weak positive association with coccidia (Fig 4.2). For eigenvector centrality however, there was a strong positive association with nematode infection status (95% credible intervals did not cross 0), and a weakly negatively association (slight crossing of 95% CIs with 0) for coccidia infection status (Fig. 4.2). There was no evidence of an association with eigenvector centrality for either bTB or Brucellosis (Fig 4.2).

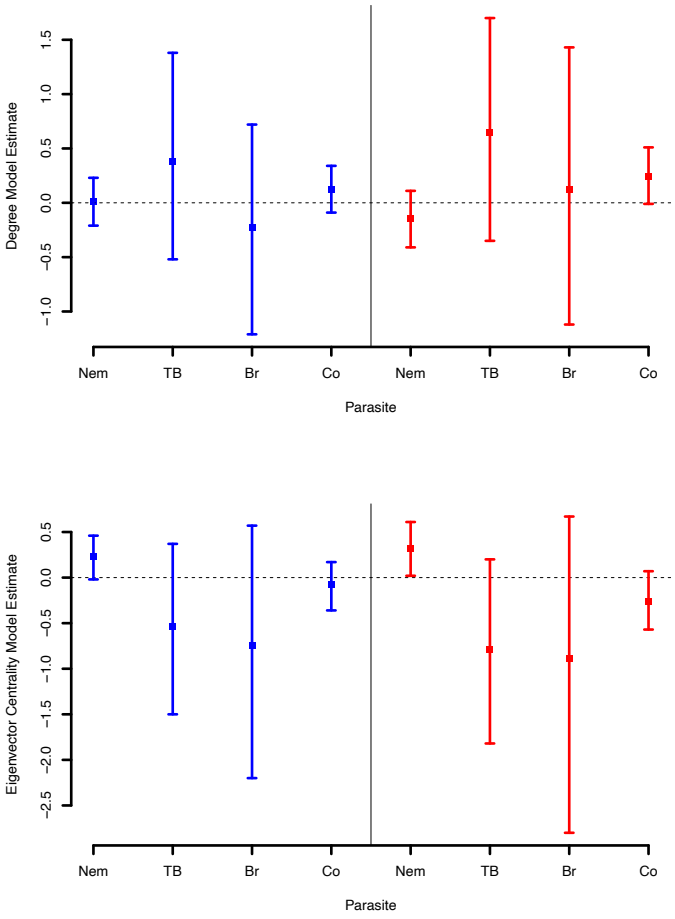

Figure 4.2 Blue, where degree or eigenvector centrality have been considered alone, red where they have been considered in the same model as each other.

Models were run with either standardised degree, standardised eigenvector centrality, or both. Full model outputs including the effects of controlling parameters are shown in Appendix 2. There appears to be no notable difference between the complete models or the individual models, with the exception of a slightly more defined posisitve relationship between degree and Coccidia infection status (Fig 4.2).

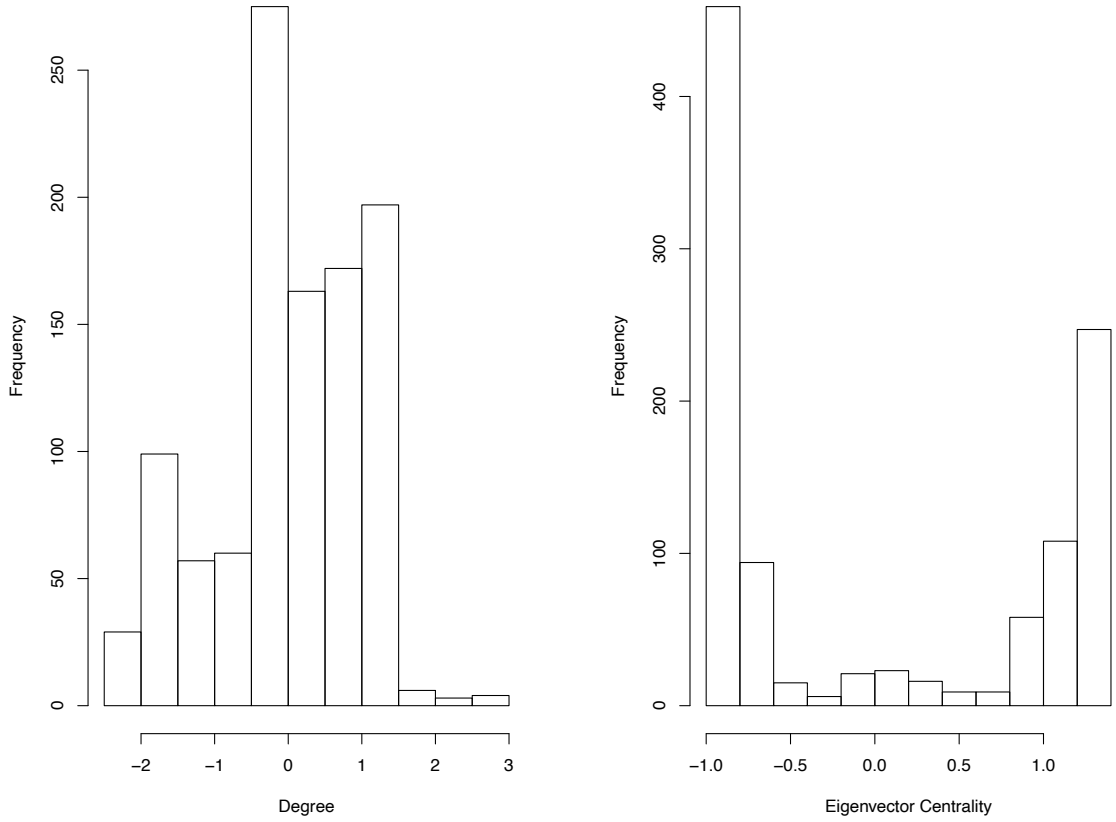

Figure 4.3 Histogram of standardised degree and of standardised eigenvector centrality

4.4.3 "Neighbourhood" analysis of local infection and coinfection risk

Neighbourhood analysis was first conducted to determine whether the prevalence of parasites at increasing distances across the network from a focal host were associated with an increase or decrease in the probability of the host being infected with that parasite or not. For bTB, neighbourhood prevalence was a strong positive predictor of infection up to network distances of 4 edges (Fig 4.4). Hence, there were clusters of bTB infection among individuals up to 4 degrees of separation from each other. For

strongyle nematodes, neighbourhood prevalence was a weak positive predictor of individual infection risk across all network distances (Fig 4.4). Coccidia neighbourhood prevalence was a strong positive predictor up to network distances of 6, but the effect then switched to being negative at distances of 7 (Fig 4.4). There was no effect of prevalence on Brucellosis at any scale (Fig 4.4).

In terms of coinfection effects, bTB neighbourhood prevalence strongly negatively predicted strongyle nematode infection status up to distance 3, but weakly negatively from 4 to 6 network distances away (Fig 4.4). There was no effect of bTB prevalence on Brucellosis or coccidia at any distance. Strongyle nematode neighbourhood prevalence did not affect coccidia infection risk up to distance 6 in the network, however they were strongly positively associated at distance 7 (Fig 4.4). Strongyle nematode prevalence did not predict either bTB or Brucellosis at any network distance, and Brucellosis prevalence did not predict the presence of any parasite in this study at any distance in the network (Fig 4.4). Similarly, coccidia neighbourhood prevalence did not predict strongyle nematode or brucellosis infection status at any distance in the network. However, coccidia prevalence weakly positively predicted bTB at distances 2 to 4 in the network, but not at any other distance (Fig 4.4).

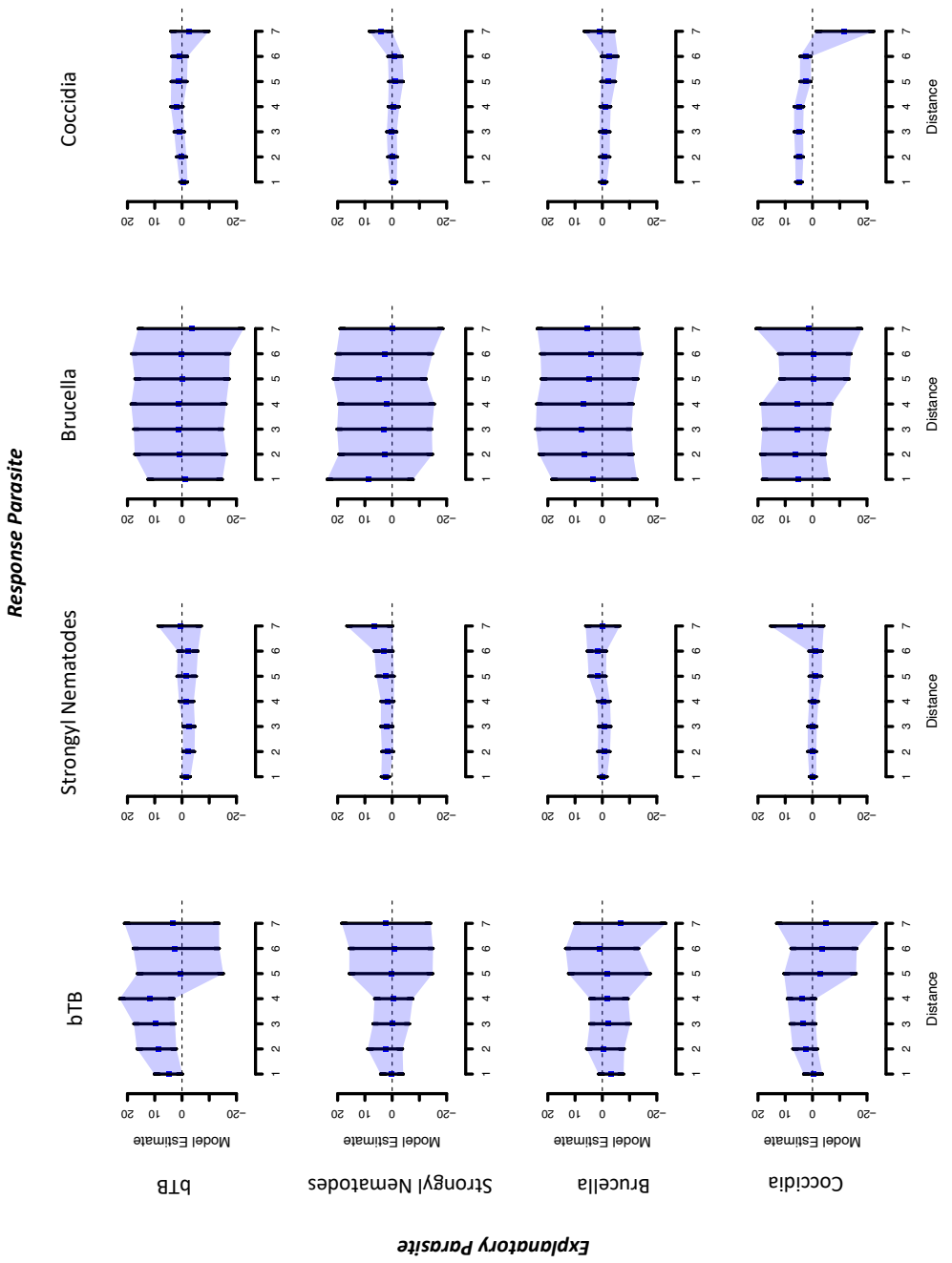

Figure 4.4 Neighbourhood analysis plots for bTB, Strongyl Nematodes, *Brucella* and Coccidia. Diagonals represent the effect of the parasite prevalence on itself.

4.4.4 Null model Test

Data were simulated as described in Appendix 1. for the purpose of testing the null model of this technique – that is a model with no coinfection interaction. As can be seen in Table 4.1, there emerges only one spurious interaction detected providing

support for the interactions detected as shown in Fig 4.4, showing considerably more interactions. These data are presented as a table and not a figure due to high variability in credible intervals seen in 50% and 100% sampling.

Table 4.1 Tabulated neighbourhood analysis results. Sampling 10%, 50% and 100% of simulated networks of 1000 individuals.

Sampling	Distance	Low CI	Mean	High CI
10%	1	-2.4	-0.9	0.6
	2	-2.4	-0.7	1
	3	-3.1	-1.1	0.8
	4	-3.4	-1.2	0.9
	5	-3.4	-1.2	0.8
Sampling	Distance	Low CI	Mean	High CI
50%	1	-729.6	-29.5	687.1
	2	-3.1	-1.2	0.7
	3	-9.4	-4.8	-0.3
	4	-19.8	-7.5	4.6
	5	-80.3	-32.1	14.8
	6	-549.7	-260.1	23.2
	7	-931.4	-121.8	681.4
Sampling	Distance	Low CI	Mean	High CI
100%	1	-743	-30.8	647.1
	2	-3	-2.1	0.6
	3	-9.2	-4.8	-0.5
	4	-19.7	-7.5	4.3
	5	-78.4	-32.1	15.1
	6	-561.3	-260.1	28.8
	7	-911	-121.8	683.3

4.5 Discussion

4.5.1 Individual-level network characteristics

Using network analysis as a proxy for understanding the contact structure between hosts, I identified eigenvector centrality as an individual-level network metric determining whether individuals are infected or not by nematodes within these African buffalo herds. Eigenvector centrality is a metric by which the influence of a node in a network is calculated with consideration to its connection to other nodes and their relative importance; a node (individual) connected to other nodes which themselves are considered important are ranked higher than nodes which are connected to less important nodes. As such, this could be considered a more holistic metric than the standard measure of the degree of a given node, which would rank higher for a node with many connections, even if these connections were themselves to nodes with little network influence. In the case of infectious disease transmission, it could be considered that connections to important neighbours with respect to their network position is in fact more important than simply than frequency of contacts. It is therefore notable that for two parasites in this network, nematodes and coccidia, there were strong effects of eigenvector centrality, whereas the effects of degree were limited small for all parasites with the exception of coccidia (Fig 4.2).

It is surprising though, that for no parasites did degree have an outright effect (the 95% Cis did not cross 0). The biological rationale for investigating degree is that it appeared to be a sensible proxy for density dependent transmission. A higher degree indicates that an individual is in contact with more neighbours. It is possible that none of these parasites are governed by density dependence alone, or, as noted above, that degree

only tells part of the story, and that eigenvector centrality is a more robust measurement of the importance of an individual in the social network.

The strong positive effect of eigenvector centrality on an individual's infection risk with nematodes is possibly an indicator that this metric is a good proxy for shared space usage. Nematodes are transmitted by eggs in the soil, and as such sharing space with more infected individuals may lead to an increase in risk of infection.

4.5.2 Neighbourhood analysis

Neighbourhood analysis has provided an insight into the effects of parasite prevalence on infection status at intermediate scales, and some insight into how parasite-parasite interactions may scale. For bTB, strongyle nematodes and coccidia, I showed that locally (i.e. sub-whole population) high prevalence is indicative of a higher risk of infection. This follows logically, that higher bTB prevalence up to an intermediate scale poses a risk, in that if an individual's neighbours have neighbours that have bTB, then the focal individual is at risk – it is not one's neighbours that pose a risk, but the neighbours of one's neighbours. For strongyle nematodes and coccidia it also follows logically that there is a generally positive trend to higher distances, as these parasites are environmentally transmitted. Environmentally transmitted parasites that have long periods of remaining infective in the environment are expected to be more ubiquitously spread in line with the distribution of the host.

The network is only partially known, representing a fraction of the animals in each herd. This does present some limitation, however simulation analyses were conducted (Table 4.1) to demonstrate that spurious interactions do not emerge regularly and that I can be confident in the reliability of the presented results.

These four parasites were chosen to focus these analyses on as there had been documented interactions between various combinations of them. Strongyle nematode prevalence was only found to associate with coccidia infection status at the largest scales. The reciprocal effect of strongyle nematodes on bTB was not detected here and there may be several reasons. It is possible that the analysis better captures bTB transmission, which is primarily a close contact transmitted parasite. This is evidenced in the single infection neighbourhood analysis of the social networks, with the notably strong effect of bTB prevalence on bTB infection probability.

There were no detectable effects of Brucella on other parasite species. This is notable given that there is the previously characterised association with bTB. However, given that the Brucella infection risk neighbourhood analysis (Fig 4.4) shows no effect of local prevalence on itself, this may in fact be unsurprising and be a way of determining the sensitivity of the technique.

Finally, one novel association that has emerged is that of the effect of coccidia prevalence on bTB. Although weak, it does appear that there is a positive predictor of bTB by coccidia prevalence up to intermediate scales. There is currently work ongoing investigating bTB-coccidia associations, and this neighbourhood analysis result is an exciting first step in determining whether there is an interaction between these parasites.

4.5.3 Future directions

bTB persistence in Kruger National Park has been experimentally determined to last approximately 4 weeks in faeces, 6 weeks from tissue samples and as low as 5 days in buried samples (Tanner and Michel, 1999). These are relatively short periods of time in

contrast to other infectious particles, such as nematode eggs. Given that these data were pooled over a whole season (6 months) for calculating network metrics there is still a significant amount of within-season behaviour that will be lost, specifically relating to mating and raising of young. Going forward, if the neighbourhood approach is to be reapplied, a more stringent temporal weighting may provide further insight into the potential interactions at these intermediate scales. While bTB transmission appears relatively well characterised in this study, a finer temporal resolution may produce better characterised networks, and as such a better spatial proxy for contacts with potential for transmission.

An area where future investigation should focus is to identify males in the system. Mating behaviour is likely to lead to differing contact structures at certain times of the year. Given the dataset comprises exclusively females, I was unable to consider these within-season behavioural variations and these data would strengthen my understanding of how contacts vary over time. Additionally, given the transmission biology of Brucella, and that aborted reproductive material can act as a means of transmission this would increase understanding of how network positioning impacts disease risk in the specific context of reproductive contacts.

While perhaps less feasible, it would perhaps be most informative to capture whole herd data. These data would be an ideal resource for understanding how network structures relate to disease metrics, however it is important to note the serious difficulty capturing such data of herds of 1000+ animals (Ezenwa and Jolles, 2015). In the absence of such detailed data capture, studies such as presented in this chapter provide sufficient insight into the nature of how host interaction heterogeneity, and interactions across a range of spatial scales affect parasite transmission. Additionally, this chapter

presents a means of detecting novel parasite-parasite interactions otherwise undetectable at previously investigated spatial scales, opening up investigation of new parasite pairs.

5 SUMMARY AND CONCLUSION

Each chapter of this thesis has tackled a key issue in spatial ecology and applied it to the understanding of infectious disease dynamics; spatial scale. This is a concept that, through overuse of the term 'scale' has a confusing past in the literature. It is hoped that the work in this thesis, by combining the development and application of novel analytical methods to two comprehensive datasets comprising mammals with different social behaviours, and multiple parasite species exhibiting a range of transmission modes and life-histories, will clarify and extend concepts relating to spatial scale in epidemiology. This final chapter summarises the key findings of the previous chapters, emphasising the novel aspects of the work presented, and considers various key, overriding themes that emerge from the separate studies, and their relevance for our understanding of the spatial dynamics of host-parasite systems more generally.

5.1.1 Thesis Summary and Novelty

In Chapter 2 I investigated the spatial clustering of a range of micro- and macro-parasites of the wood mouse. This chapter is novel in that this is the first cross-parasite species comparison of clustering by use of the K function and associated statistical tests that I am aware of. These parasite species ranged in transmission mode (close contact, environmentally via soil and flea borne). I found a difference in spatial clustering between WMHV cases, a close contact transmitted parasite, and that of the host and other parasites. However, I found no notable differences between the host and environmentally and flea borne parasites.

In Chapter 3, I sought out to understand how parasite-parasite interactions that occur within the host, have consequences for transmission and infection risk at spatial scales beyond the individual host. To achieve this I developed a new analytical method, neighbourhood analysis, that quantifies the infection prevalence of parasites over incremental spatial neighbourhoods around each host, and then relates that prevalence to the infection status of that parasite, or other parasites, within focal hosts. This is the first attempt to empirically determine the spatial scale over which within-host interspecific parasite interactions extend. The method was tested using the known antagonistic interaction between *H. polygyrus* and *E. hungaryensis*, and the known non-interaction between *H. polygyrus* and *E. apionodes* (Knowles *et al.*, 2013), and showed that parasite-parasite interactions can be detected from among-host data, but only by examining those data at an appropriate spatial scale. The method was then applied to investigate all possible pairs of parasites in the system, which revealed a number of novel interactions not previously detected in this study system, highlighting its potential for uncovering potential within-host interactions not detectable by other means.

In Chapter 4, I sought to extend the question of the spatial (or social) scale of interspecific parasite interaction effects, using a different study system (African buffaloes) with different social behaviour from territorial wood mice; herding behaviour. This involved adapting the neighbourhood analysis approach from Chapter 3 and allowing the exploration of herding animals and determining how both infection risk and coinfection interactions scale in these types of systems. This chapter first showed that individual infection risk was most associated with the eignenvector centrality of the individual; suggesting that risk isn't purely due to the number of contacts an individual, but due to the number of secondary contacts each of those contacted animals have. Applying this method to the coinfection data then led to the

discovery of a novel parasite-parasite interaction, of coccidial on bTB infections, potentially furthering our understanding of the transmission dynamics of this important pathogen.

5.1.2 Comparative clustering – does transmission mode play a role in the spatial distribution of cases?

The analyses presented in Chapter 2 provide tentative support for the hypotheses that differences in parasite transmission biology may leave different signals on the observed degree of parasite clustering relative to that of the host at different spatial scales.

I found evidence that WMHV exhibits a higher degree of clustering than the host, as well as other parasites, both environmentally transmitted (*E. hungaryensis* and *H. polygyrus*) and flea-borne (*B. taylorii* and *T. grosi*) at scales greater than 20 metres. WMHV is the only parasite I analysed to be transmitted via close contact of individuals, but this increased clustering relative to that of its host, is similar to what Carslake *et al.* (2005) observed for another close contact transmitting virus, cowpox. While it is difficult to generalise too far beyond two examples, the similarities in these responses for these similarly transmitted parasites do suggest contact transmitted viruses tend to show higher levels of spatial clustering than their host, resulting in higher levels of spatial aggregation of cases across the landscape. However, it must be remembered that the K function is a cumulative measure; whilst we may observe this increased relative clustering over 20 metres, processes operating at smaller spatial scales may be impacting this.

I found no evidence that environmentally transmitted parasites in this dataset (the nematode *H. polygyrus*, or the gut-dwelling *Eimeria* spp.) show significantly different clustering from the host animals over any spatial scales. This contrasts with what Snow found in the cholera cases in London, mentioned in the Chapter 2, which is also environmentally transmitted. However, cholera was not ubiquitous in the environment, and the clustering was detected around sources of cholera (water pumps) as opposed to cases, potentially explaining the differences observed here. Wood mice defecate in the environment throughout their home range. This would imply that *H. polygyrus* eggs and *Eimeria spp.* oocysts are also deposited throughout their range. Given that *H. polygyrus* eggs and *Eimeria spp.* oocysts can remain viable for some time in the soil (likely several months), the environment which they occupy is likely to pose an infection risk to others for a long period of time. Hence the deposition of parasite eggs or oocysts into the environment from an infected animal, and the subsequent uptake and infection of them into another animal, will act to decouple the observed occurrence of cases. As such it is then not surprising that we see the distribution of parasites to be not different from that of the hosts.

For flea borne parasites, I had hypothesised that as fleas are able to disperse independently of the host, that I may see clustering operate at potentially larger scales than the host home range. Surprisingly, *B. grahamii* and *B. taylorii* do not differ from host clustering at any spatial scale. This leads to two potential conclusions. Either, flea vectors are spatially tied significantly closer to the host than I had expected, or the scale of analysis is insufficient to detect any difference in clustering.

Spatiotemporal analysis is the direction that future analyses of these types should take. Given my analyses have been undertaken on a legacy dataset (i.e., one collected before

the present analyses were conceived), I was unable to direct sampling with respect to a spatiotemporal analysis. The benefits to the spatiotemporal approach as undertaken by Carslake *et al.*, (2005) is that the clustering of cases in both space and time can be given due consideration. The key difference in the dataset they analysed and the one that I have analysed is that there is a significant gap in trapping each year (from January – May) in the data examined here, as opposed to full annual coverage across multiple years. Hence I was not able to examining temporal clustering of cases over extended windows of time, preventing assessment of the temporal extent of prior space use by one animal on subsequent infection risk by another. Such temporal effects may be important, for example, in distinguishing *H. polygyrus* clustering from that of *E. hungaryensis*, if one species has more long-lived infective stages in the environment than the other.

5.1.3 Neighbourhood analysis as a tool to understand coinfection interactions

I have been able to demonstrate, using a well-characterised parasite-parasite interaction in a wild mammal population (Knowles *et al.*, 2013), that there are detectable consequences of these within-host interactions for between-host transmission. Given this suppressive interaction of *H. polygyrus* on *E. hungaryensis*, and the reduction in *E. hungaryensis* oocysts shed by *H. polygyrus* coinfected hosts (Knowles *et al.*, 2013), it follows logically that the transmission potential of *E. hungaryensis* should be severely decreased. I show that this within-host interaction does indeed suppress *E. hungaryensis* transmission, and that the signal of this negative interaction is detectable up to two home ranges around the host, with strongest effects closest to a single home range.

This validation of the method has allowed a much more thorough exploration of other potential parasite interactions in this system. I have detected 6 novel interactions at

scales beyond the host that have not previously been reported in this system. I found that the neighbourhood prevalence of *H. polygyrus* affected individual infection risk of *Bartonella spp.* positively and that of *T. grosi* negatively. However, despite the contrasting directions of these effects, the peaks of where the spatial signal was most strong was the same, between 1 and 2 home ranges. Both parasites are flea transmitted, suggesting that there may be a link between transmission mode and the spatial scale over which the signal of any interactive effects occur.

A further implication of these analyses is that the spatial scale of observation is vital in detecting these interactions, given the seemingly local nature of many that I have detected. This technique therefore, represents a significant move forward in the detection of coinfection interactions from observational data. Larger scale, population-based analyses of coinfection interactions have previously failed to detect any significant association between parasites from cross-sectional data because of this sensitivity of detection to the spatial scale examined (Fenton *et al.*, 2014). The results presented here suggest that examining whole-population data effectively averages across all spatial scales, thereby obscuring the local processes of transmission interference arising from within-host interactions that play out over much smaller spatial scales around each individual host.

This form of rodent mark-recapture experiment on regularly spaced grids is common to many studies of wildlife, with similarly collected data (Turner *et al.*, 2014). Given the utility of the technique, and the pre-existing data available, application of neighbourhood analysis to both new and legacy data of this form may provide significantly more insight into the interactions between parasites, and therefore their community ecology.

Overall I have shown that within-host interactions between coinfecting parasites can affect their localised transmission dynamics, leaving a signal of that interaction at spatial scales beyond the individual host. Furthermore the spatial extent of that effect operates across different spatial scales for different parasites, and is likely to reflect their spatial scale of transmission. There are many human, livestock and wildlife disease systems where there are well established within-host parasite interactions among coinfecting parasites, and a growing body of research has examined the consequences of those interactions for the success and impact (beneficial or detrimental) of disease treatment approaches on individual host health (Griffiths *et al.*, 2011, 2015). However, my results suggest that there could be knock-on, between-host consequences of such treatments, particularly in communities experiencing high coverage mass drug administration, for localised transmission dynamics of non-target parasites, with implications for infection risk even among non-treated individuals.

5.1.4 Extending neighbourhood analysis to social networks

Neighbourhood analysis, adapted for use on social networks (Chapter 4), has provided an insight into the effects of parasite prevalence on infection status at intermediate scales, and some insight into how parasite-parasite interactions may scale. I showed that locally (i.e. sub-whole population) high prevalence is indicative of a higher risk of infection for bTB, strongyle nematodes and coccidia.

These four parasites were chosen to focus these analyses on as there had been documented interactions between various combinations of them. Strongyle nematode prevalence was only found to associate with coccidia infection status at the largest

scales. The reciprocal effect of strongyle nematodes on bTB was not detected here and there may be several reasons. It is possible that the analysis better captures bTB transmission, which is primarily a close contact transmitted parasite. This is evidenced in the single infection neighbourhood analysis of the social networks, with the notably strong effect of bTB prevalence on bTB infection probability. There were no effects of Brucella on other parasite species detectable in these models which is notable given that there is the previously characterised association with bTB. Given that the Brucella infection risk neighbourhood analysis shows no effect of local prevalence on itself however, this may in fact be unsurprising and act as a means of determining the sensitivity of the technique. One novel association that has emerged is that of the effect of coccidia prevalence on bTB. Although weak, it does appear that there is a positive predictor of bTB by coccidia prevalence up to intermediate scales. There is currently work ongoing investigating bTB-coccidia associations, and this neighbourhood analysis result is an exciting first step in determining whether there is an interaction between these parasites.

While perhaps not practically feasible, it would perhaps be most informative to capture whole herd data. These data would be an ideal resource for understanding how neighbourhoods of animals are spread across a social network. However it is worth noting the serious difficulty capturing such data of herds of 1000+ animals (Ezenwa and Jolles, 2015). In the absence of such detailed data capture, studies such as presented in Chapter 5 provide sufficient insight into the nature of how host interaction heterogeneity, and interactions across a range of spatial scales affect parasite transmission.

5.1.5 Overall conclusion

The theme binding this thesis together is that of spatial scale - from scale of clustering, to scale of coinfection interactions. Using pre-existing (Chapters 2 and 4) and bespoke techniques (Chapter 3 and 4), I have used 2 comprehensive datasets to tackle these issues in multi-parasite systems. Spatial scale is ultimately important in understanding these systems, and in disease ecology more broadly. Future analyses should consider it from the data collection point of view.

6 Appendix 1. Non-linear adjustment factor for network connectance for use in simulation of large networks

Networks were constructed as described in Chapter 4. For the purposes of simulating networks of a larger size, but retaining key aspects of the structure of the overall network, it is appropriate to scale connectance in a non-linear fashion. Given the study system in question are buffalo, which occupy real space, it is fair to assume that as the number of individuals increases, the space they will occupy increases. I assume this will increase with the square of the number of individuals, like area increases with the square of the radius of a circle. This will ultimately have an effect on the network metrics, given that individuals further apart in space may be less likely to encounter each other. Given that area increases non-linearly, it is appropriate to use a similarly considered adjustment factor in this case.

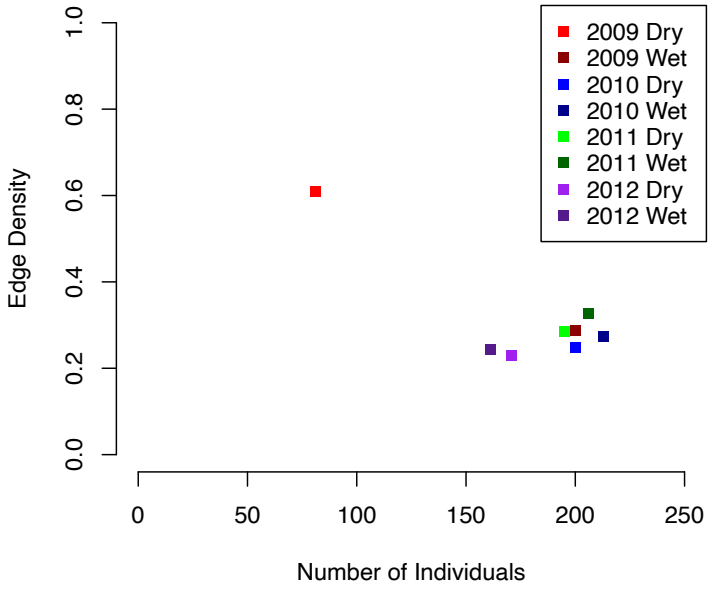

Figure A1.1 *Edge density of each network in relation to the number of individuals used to construct the network. The 2009 dry season is a clear outlier with regards to both number of individuals and edge density.*

Based on the data displayed in Fig A1.1, I have elected to exclude the network metrics from the 2009 dry season as a means of calculating an adjustment factor for the scaling up of connectance. The mean size of the network, excluding the 2009 dry season is 192.28 and the mean connectance is 0.270. Using the formula,

$$C_{target} = C_{real}\left(\frac{{N_{real}}^2}{{N_{target}}^2}\right) \quad \text{A1}$$

to calculate the target connectance, I multiply the observed connectance (0.270) by a non-linear adjustment factor $\frac{{N_{real}}^2}{{N_{target}}^2}$, which accounts for differences between the mean observed population (192.28) and those of networks to be simulated.

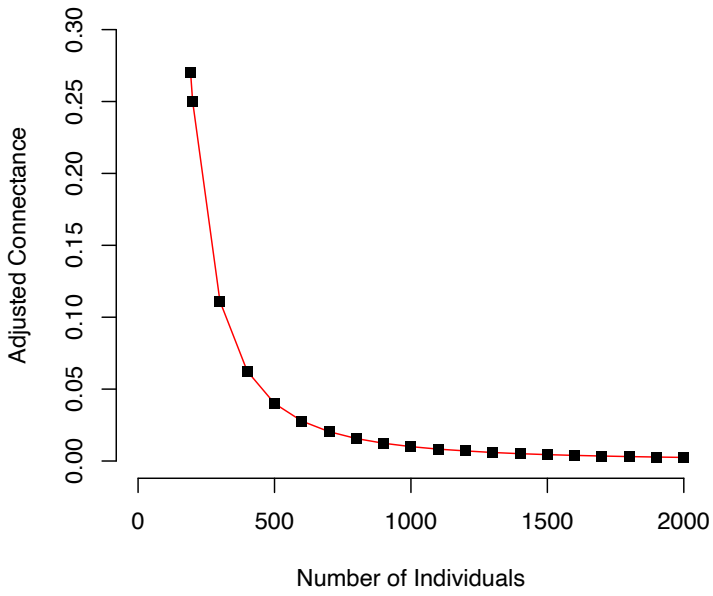

Figure A1.2 *Adjusted connectance based on application of formula A1*

Connectance is therefore calculated as shown in Fig A1.2 based on the observed values in the data. These adjusted connectances are then applied when simulating networks larger than what is seen in the observed data.

7 Appendix 2. Supplementary model outputs

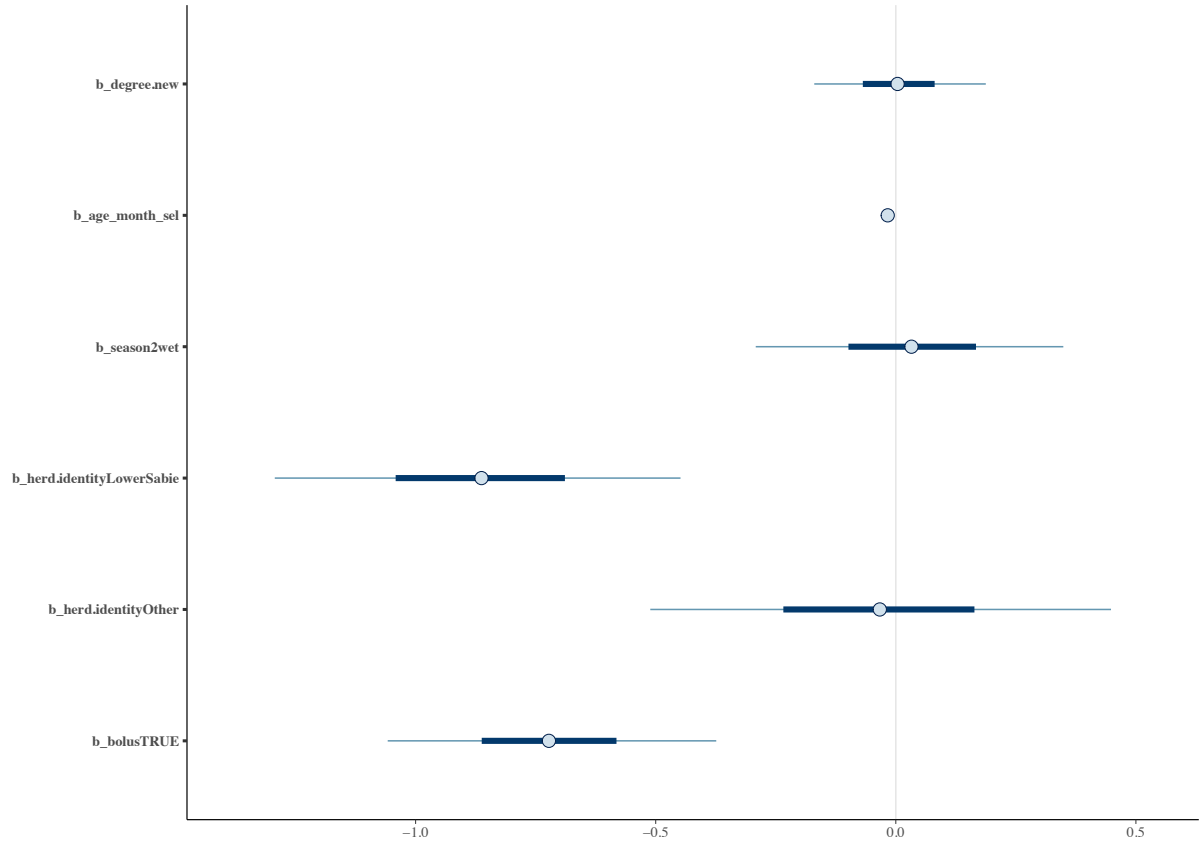

Figure A2.1 Nematode: Degree Only

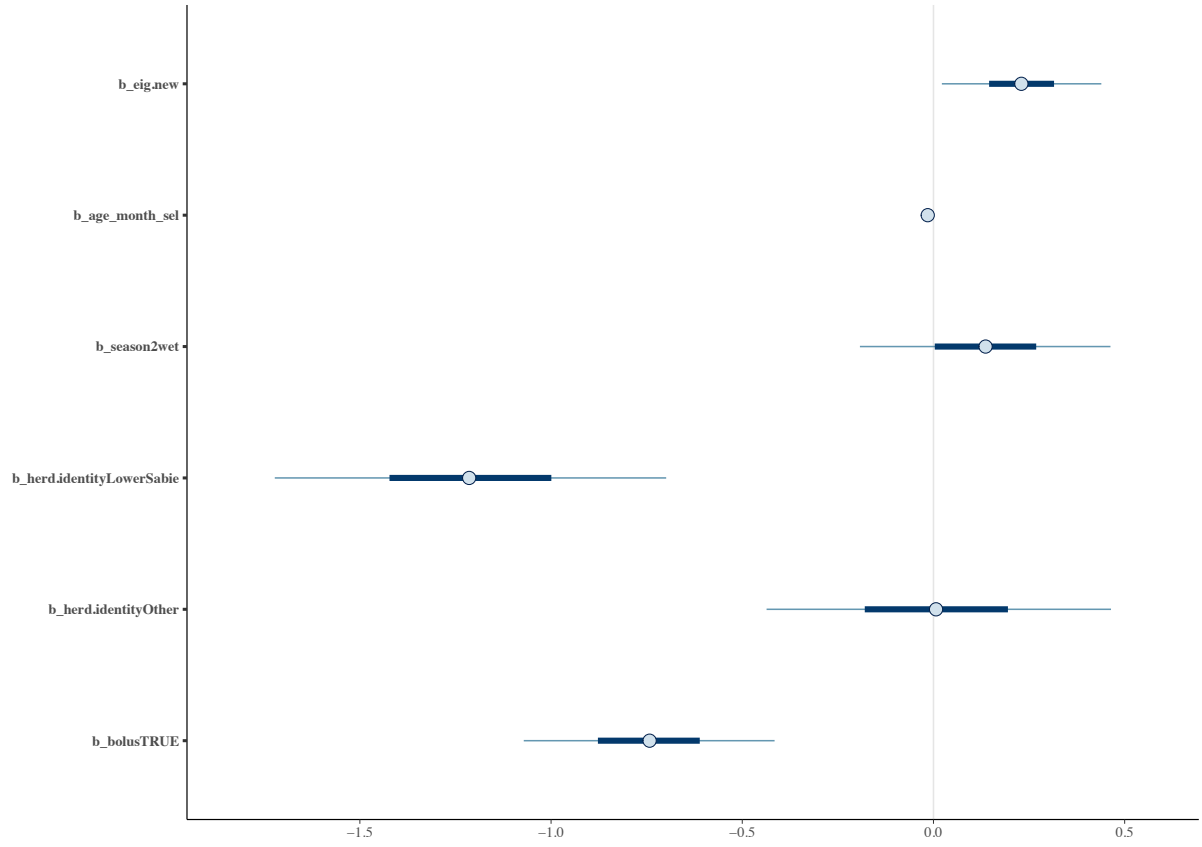

Figure A2.2 Nematode: Eigenvector Centrality Only

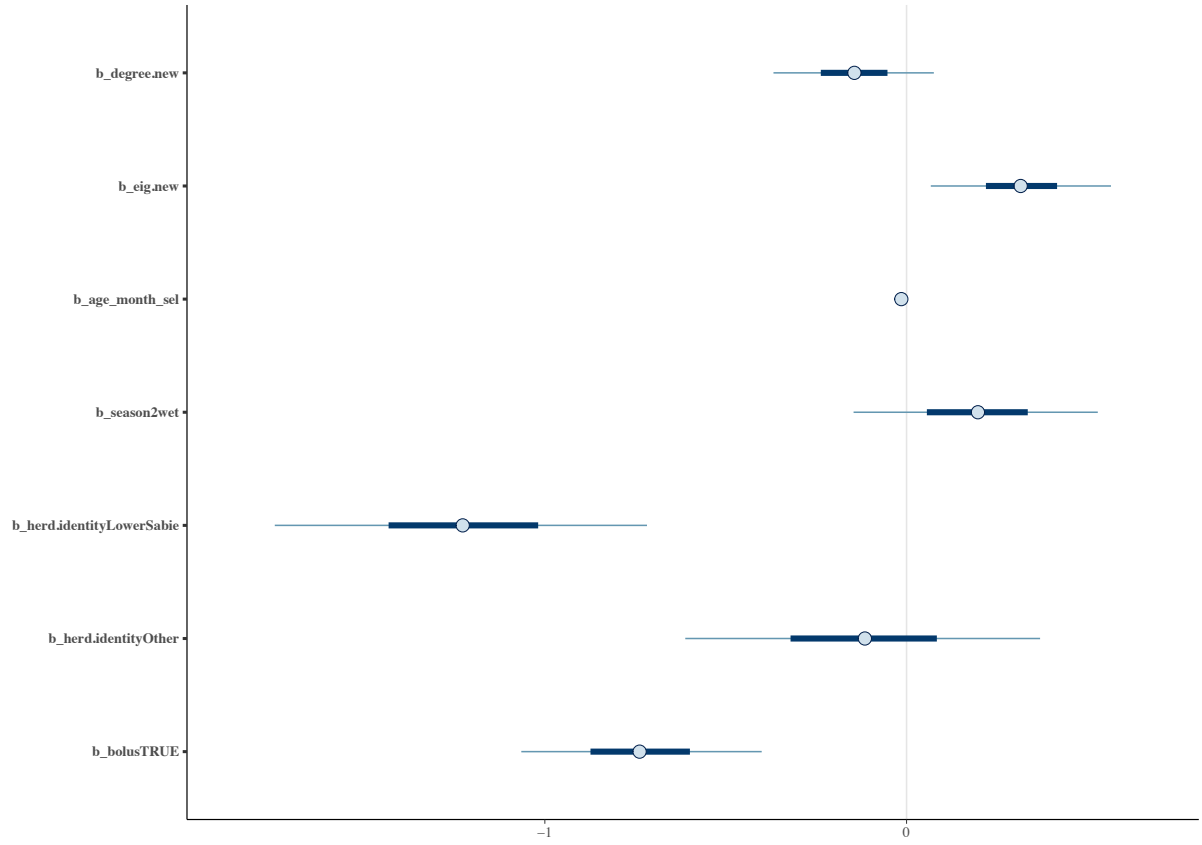

Figure A2.3 Nematode: Complete Model

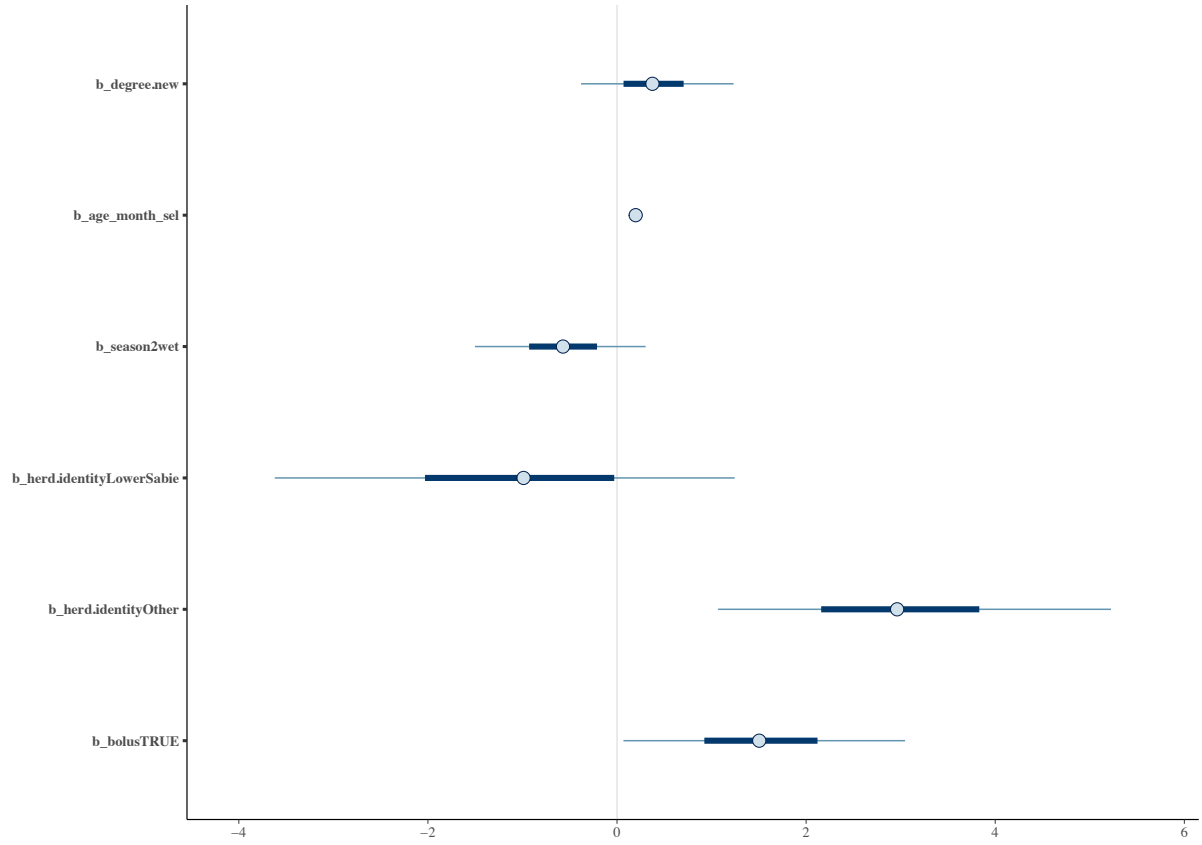

Figure A2.4 bTB: Degree Only

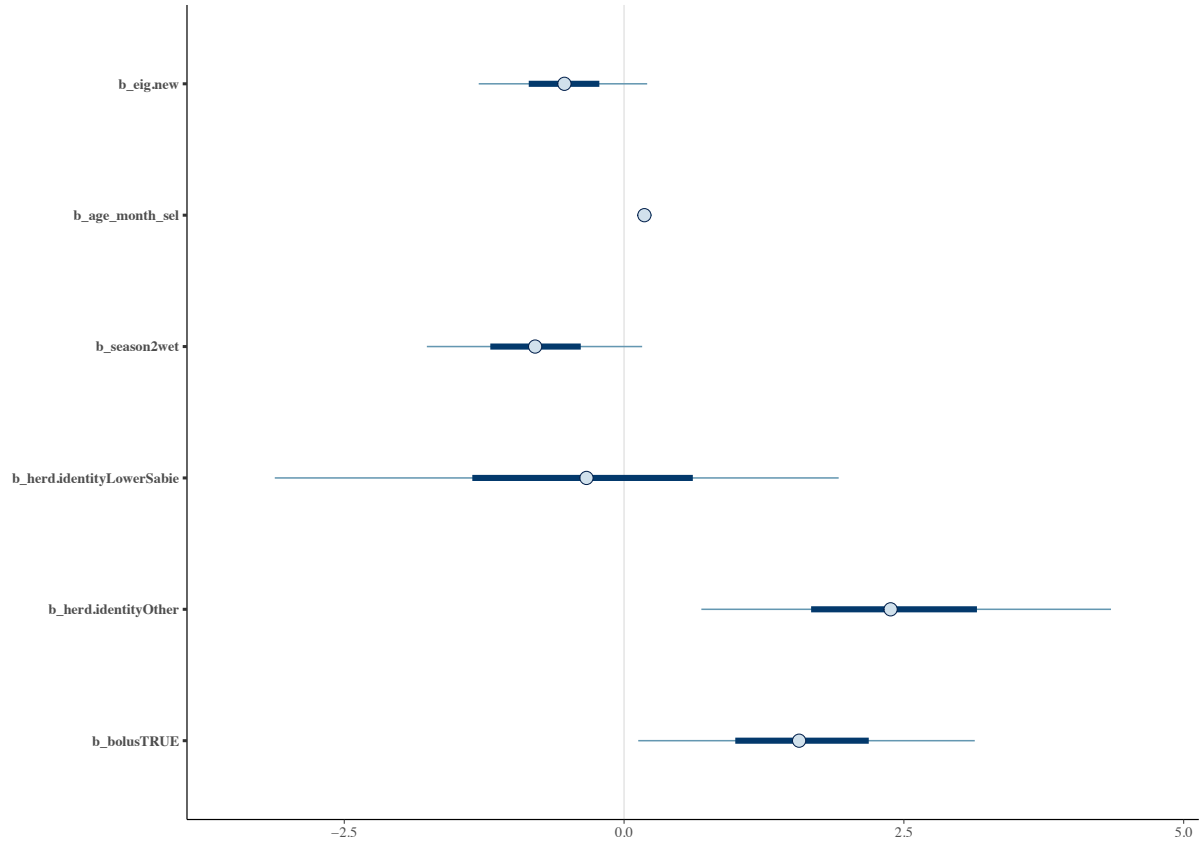

Figure A2.5 bTB: Eigenvector Centrality Only

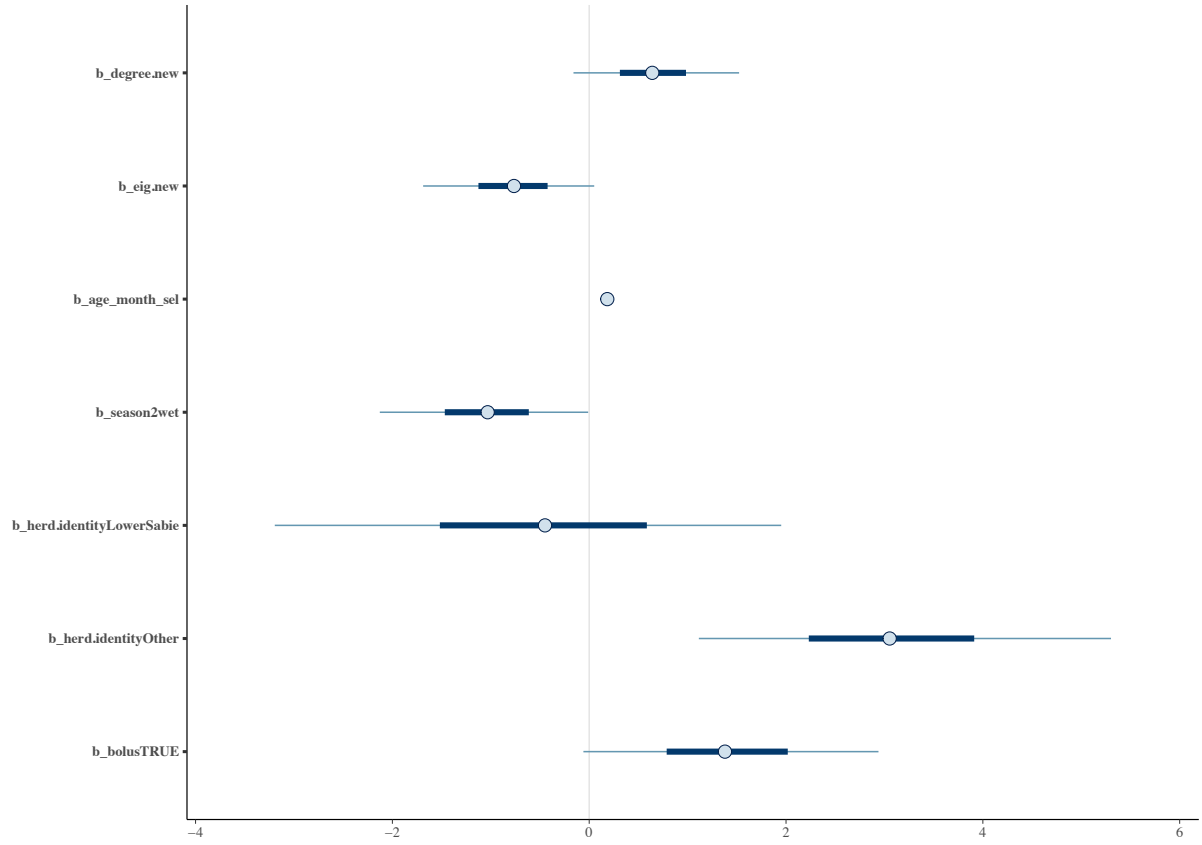

Figure A2.6 bTB: Complete Model

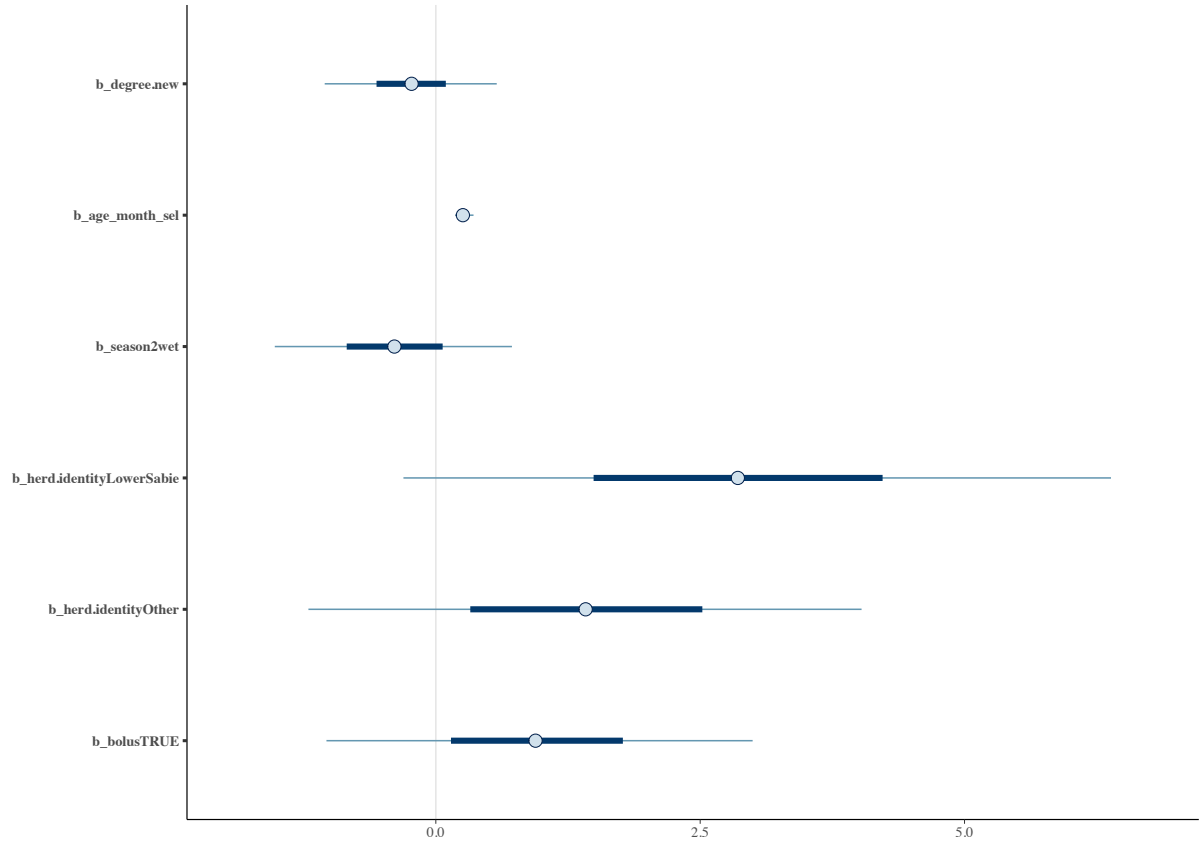

Figure A2.7 Brucella: Degree Only

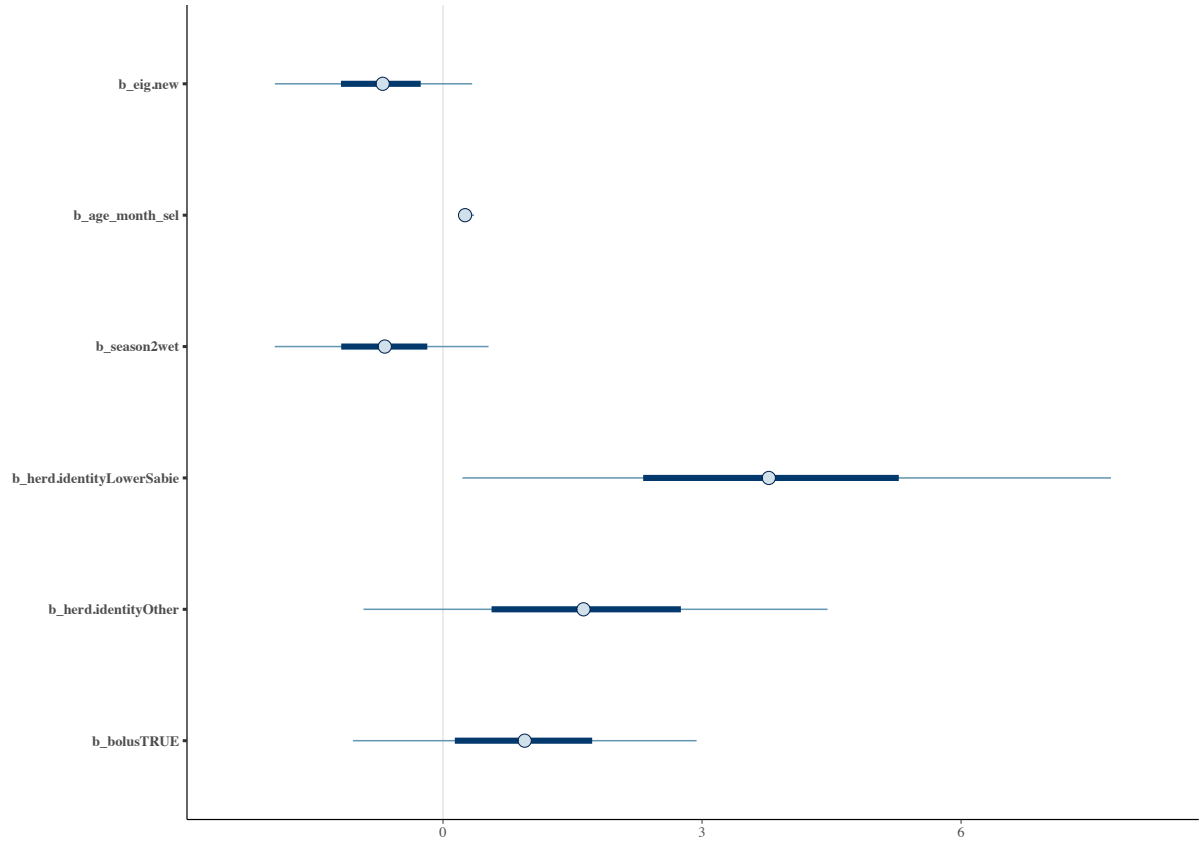

Figure A2.8 Brucella: Eigenvector Centrality Only

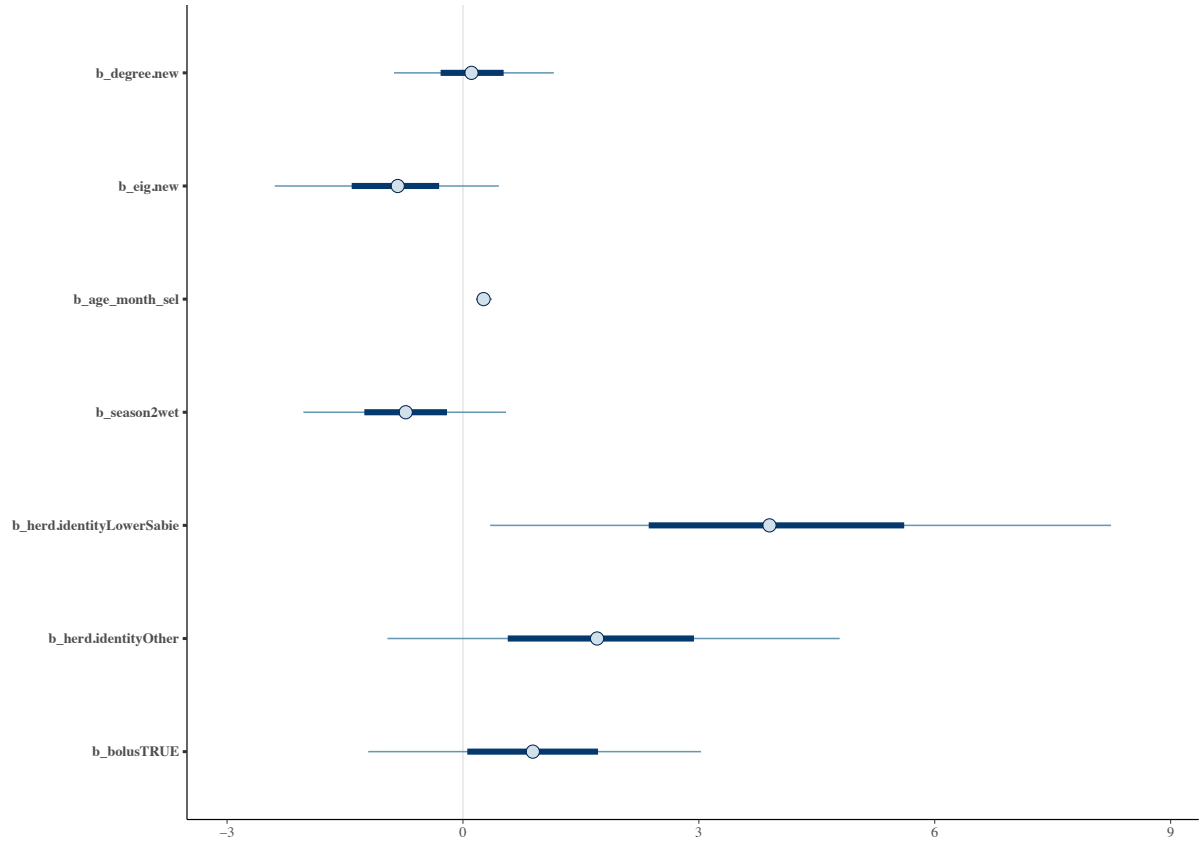

Figure A2.9 Brucella: Complete Model

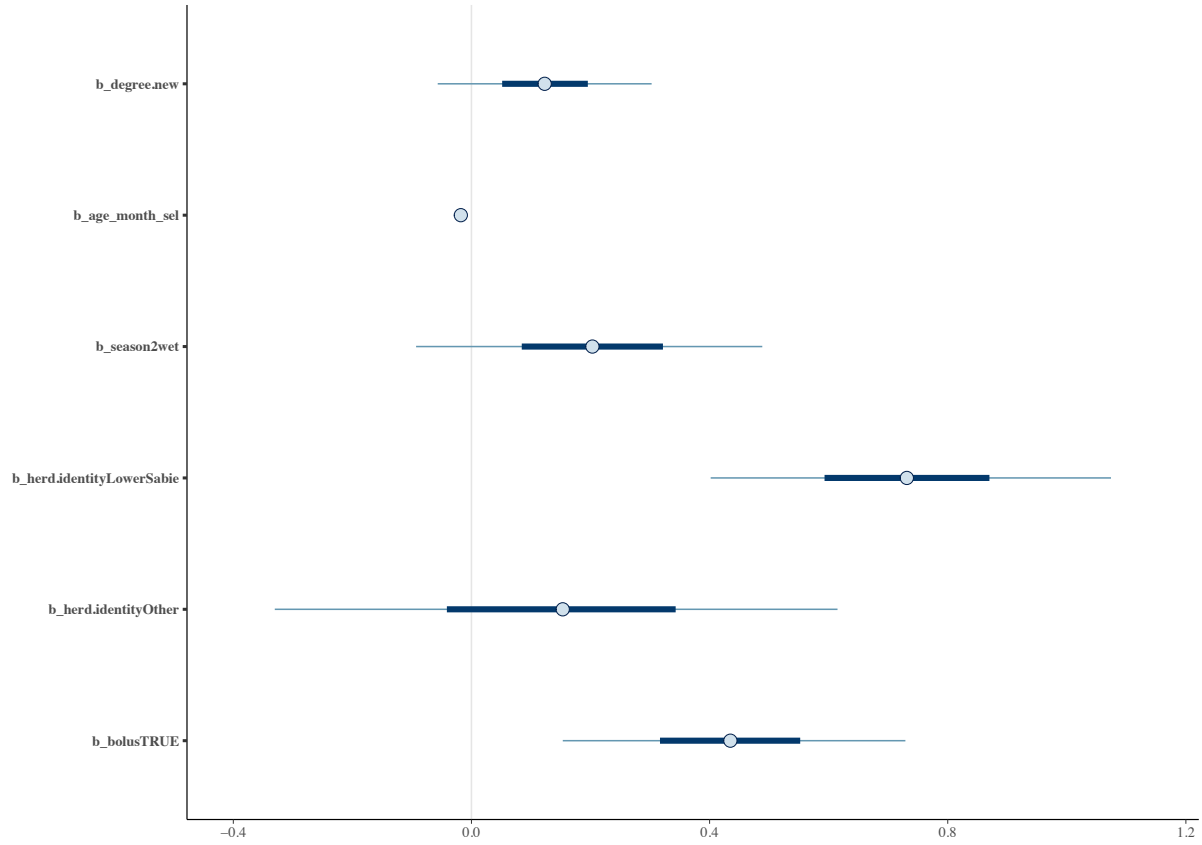

Figure A2.10 Coccidia: Degree Only

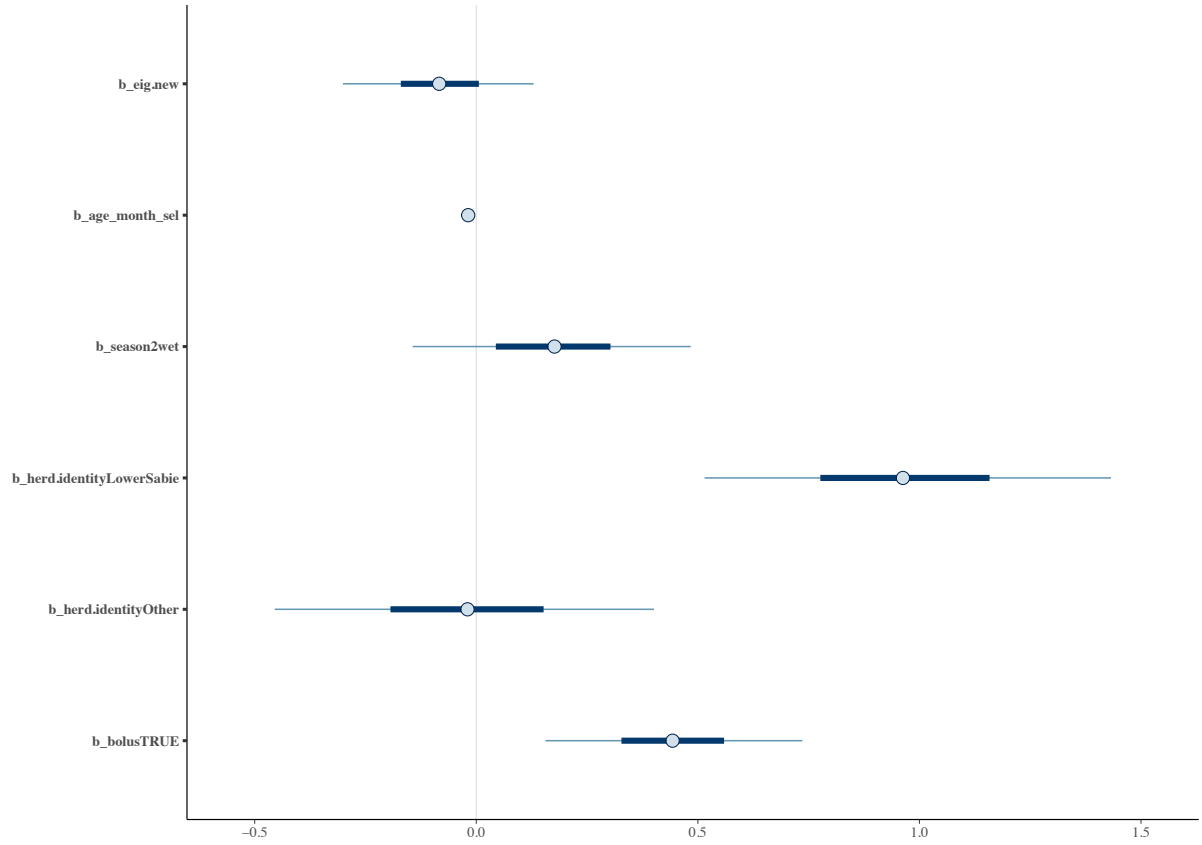

Figure A2.11 Coccidia: Eigenvector Centrality Only

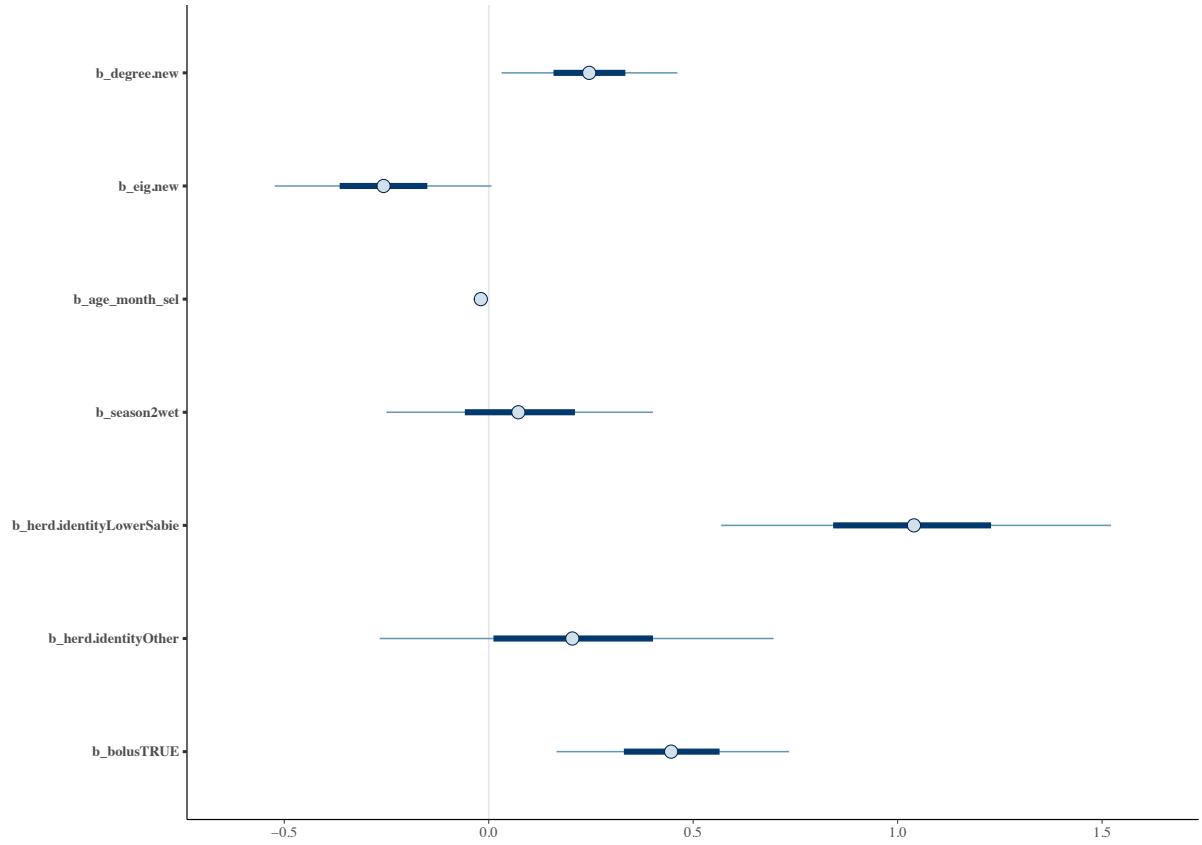

Figure A2.12 Coccidia: Complete Model

8 REFERENCES

Aiello, C. M. *et al.* (2014) 'Disease dynamics during wildlife translocations: Disruptions to the host population and potential consequences for transmission in desert tortoise contact networks', *Animal Conservation*, 17(S1), pp. 27–39. doi: 10.1111/acv.12147.

Altizer, S. *et al.* (2003) 'Social Organization and Parasite Risk in Mammals: Integrating Theory and Empirical Studies', *Annual Review of Ecology, Evolution, and Systematics*, 34(1), pp. 517–547. doi: 10.1146/annurev.ecolsys.34.030102.151725.

Anderson, R. M. and May, R. M. (1992) *Infectious Diseases of Humans Dynamics and Control*. Oxford University Press.

Antonovics, J. *et al.* (2017) 'The evolution of transmission mode', *Philosophical Transactions of the Royal Society B: Biological Sciences*, 372(1719), pp. 7–11. doi: 10.1098/rstb.2016.0083.

Babu, S. and Nutman, T. B. (2016) 'Helminth-Tuberculosis Co-infection: An Immunologic Perspective', *Trends in Immunology*. Elsevier Ltd, 37(9), pp. 597–607. doi: 10.1016/j.it.2016.07.005.

Baddeley, A., Rubak, E. and Turner, R. (2015) *Spatial Point Patterns: Methodology and Applications with R*. London: Chapman and Hall/CRC Press. Available at: http://www.crcpress.com/Spatial-Point-Patterns-

Methodology-and-Applications-with-R/Baddeley-Rubak-Turner/9781482210200/.

Bansal, S., Grenfell, B. T. and Meyers, L. A. (2007) 'When individual behaviour matters: Homogeneous and network models in epidemiology', *Journal of the Royal Society Interface*, 4(16), pp. 879–891. doi: 10.1098/rsif.2007.1100.

Begon, M. *et al.* (2002) 'A clarification of transmission terms in host-microparasite models: Numbers, densities and areas', *Epidemiology and Infection*, 129(1), pp. 147–153. doi: 10.1017/S0950268802007148.

Bielby, J. *et al.* (2014) 'Badger responses to small-scale culling may compromise targeted control of bovine tuberculosis', *Proceedings of the National Academy of Sciences*, 111(25), pp. 9193–9198. doi: 10.1073/pnas.1401503111.

Blyton, M. D. J. *et al.* (2014) 'Not all types of host contacts are equal when it comes to E. coli transmission', *Ecology Letters*, 17(8), pp. 970–978. doi: 10.1111/ele.12300.

Budischak, S. A. *et al.* (2018) 'Competing for blood: The ecology of parasite resource competition in human malaria-helminth co-infections', *Ecology Letters*, pp. 536–545. doi: 10.1111/ele.12919.

Carne, C. *et al.* (2014) 'The risk of disease to great apes: Simulating disease spread in orang-utan (pongo pygmaeus wurmbii) and chimpanzee (pan troglodytes schweinfurthii) association networks', *PLoS ONE*, 9(4). doi:

10.1371/journal.pone.0095039.

Carslake, D. *et al.* (2005) 'Space-time clustering of cowpox virus infection in wild rodent populations', *Journal of Animal Ecology*, 74(4), pp. 647–655. doi: 10.1111/j.1365-2656.2005.00966.x.

Carslake, D. *et al.* (2006) 'Inference of cowpox virus transmission rates between wild rodent host classes using space-time interaction', *Proceedings of the Royal Society B: Biological Sciences*, 273(1588), pp. 775–782. doi: 10.1098/rspb.2005.3400.

Carter, S. P. *et al.* (2007) 'Culling-induced social perturbation in Eurasian badgers Meles meles and the management of TB in cattle: An analysis of a critical problem in applied ecology', *Proceedings of the Royal Society B: Biological Sciences*, 274(1626), pp. 2769–2777. doi: 10.1098/rspb.2007.0998.

Clerc, M (2017) *The causes and consequences of within-host parasite interactions in wild wood mice*. University of Edinburgh.

Clerc, Melanie (2017) 'The causes and consequences of within-host parasite interactions in wild wood mice'.

Clerc, M. *et al.* (2019) 'Age affects antibody levels and anthelmintic treatment efficacy in a wild rodent', *International Journal for Parasitology: Parasites and Wildlife*, 8(September 2018), pp. 240–247. doi: 10.1016/j.ijppaw.2019.03.004.

Corner, L. A. L., Pfeiffer, D. U. and Morris, R. S. (2003) 'Social-network analysis of Mycobacterium bovis transmission among captive brushtail

possums (Trichosurus vulpecula)', *Preventive Veterinary Medicine*, 59(3), pp. 147–167. doi: 10.1016/S0167-5877(03)00075-8.

Cox, F. E. (2001) 'Concomitant infections, parasites and immune responses.', *Parasitology*, 122, pp. S23--S38. doi: 10.1017/S003118200001698X.

Diggle, P. J. (2013) *Statistical Analysis of Spatial and Spatio-Temporal Point Patterns*. 3rd edn. Chapman and Hall/CRC.

Dungan, J. L. *et al.* (2002) 'A balanced view of scale in spatial statistical analysis', *Ecography*, 25(February), pp. 626–640. doi: 10.1034/j.1600-0587.2002.250510.x.

Ezenwa, V. O. *et al.* (2010) 'Hidden consequences of living in a wormy world: nematode-induced immune suppression facilitates tuberculosis invasion in African buffalo.', *The American naturalist*, 176(5), pp. 613–624. doi: 10.1086/656496.

Ezenwa, V. O. and Jolles, A. E. (2015) 'Opposite effects of anthelmintic treatment on microbial infection at individual versus population scales', *Science*, 347(6218), pp. 175–177. doi: 10.1126/science.1261714.

Fenton, A. *et al.* (2002) 'Parasite transmission: Reconciling theory and reality', *Journal of Animal Ecology*, 71(5), pp. 893–905. doi: 10.1046/j.1365-2656.2002.00656.x.

Fenton, A. (2008) 'Worms and germs: the population dynamic consequences of microparasite-macroparasite co-infection.', *Parasitology*,

135(13), pp. 1545–1560. doi: 10.1017/S003118200700025X.

Fenton, A. (2013) 'Dances with worms: the ecological and evolutionary impacts of deworming on coinfecting pathogens', *Parasitology*, 140(09), pp. 1119–1132. doi: 10.1017/S0031182013000590.

Fenton, A. *et al.* (2014) 'The reliability of observational approaches for detecting interspecific parasite interactions: Comparison with experimental results', *International Journal for Parasitology*. Australian Society for Parasitology Inc., 44(7), pp. 437–445. doi: 10.1016/j.ijpara.2014.03.001.

Fortin, M. and Dale, M. R. T. (2005) *Spatial Analysis A Guide for Ecologists*. Cambridge University Press. doi: https://doi-org.liverpool.idm.oclc.org/10.1017/CBO9780511542039.

Gatrell, A. C. *et al.* (1996) 'Spatial Point Pattern Analysis and Its Application in Geographical Epidemiology', *Transactions of the Institute of British Geographers*, 21(1), pp. 256–274. Available at: http://www.jstor.org/stable/622936.

Godfrey, S. S. *et al.* (2009) 'Network structure and parasite transmission in a group living lizard, the gidgee skink, Egernia stokesii', *Behavioral Ecology and Sociobiology*, 63(7), pp. 1045–1056. doi: 10.1007/s00265-009-0730-9.

Godfrey, S. S. *et al.* (2010) 'Social network structure and parasite infection patterns in a territorial reptile, the tuatara (Sphenodon punctatus)', *International Journal for Parasitology*. Australian Society for Parasitology Inc., 40(13), pp. 1575–1585. doi: 10.1016/j.ijpara.2010.06.002.

Gorsich, E. E. *et al.* (2018a) 'Opposite outcomes of coinfection at individual and population scales', *Proceedings of the National Academy of Sciences*, 115(29), p. 201801095. doi: 10.1073/pnas.1801095115.

Gorsich, E. E. *et al.* (2018b) 'Opposite outcomes of coinfection at individual and population scales', *Proceedings of the National Academy of Sciences*, 115(29), pp. 7545–7550. doi: 10.1073/pnas.1801095115.

Gorsich, E. E., Ezenwa, V. O. and Jolles, A. E. (2014) 'Nematode-coccidia parasite co-infections in African buffalo: Epidemiology and associations with host condition and pregnancy', *International Journal for Parasitology: Parasites and Wildlife*. Australian Society for Parasitology, 3(2), pp. 124–134. doi: 10.1016/j.ijppaw.2014.05.003.

Grabowski, M. K. *et al.* (2014) 'The Role of Viral Introductions in Sustaining Community-Based HIV Epidemics in Rural Uganda: Evidence from Spatial Clustering, Phylogenetics, and Egocentric Transmission Models', *PLoS Medicine*, 11(3). doi: 10.1371/journal.pmed.1001610.

Graham, C. S. *et al.* (2001) 'Influence of human immunodeficiency virus infection on the course of hepatitis C virus infection: A meta-analysis', *Clinical Infectious Diseases*, 33(4), pp. 562–569. doi: 10.1086/321909.

Griffin, R. H. and Nunn, C. L. (2012) 'Community structure and the spread of infectious disease in primate social networks', *Evolutionary Ecology*, 26(4), pp. 779–800. doi: 10.1007/s10682-011-9526-2.

Griffiths, E. C. *et al.* (2011) 'The nature and consequences of coinfection in

humans', *Journal of Infection*. Elsevier Ltd, 63(3), pp. 200–206. doi: 10.1016/j.jinf.2011.06.005.

Griffiths, E. C. *et al.* (2015) 'Reported co-infection deaths are more common in early adulthood and among similar infections', *BMC Infectious Diseases*. BMC Infectious Diseases, 15(1), pp. 1–7. doi: 10.1186/s12879-015-1118-2.

Hahn, U. (2012) 'A studentized permutation test for the comparison of spatial point patterns', *Journal of the American Statistical Association*, 107(498), pp. 754–764. doi: 10.1080/01621459.2012.688463.

Hajnická, V. *et al.* (2017) 'Tick-Borne Transmission of Murine Gammaherpesvirus 68', *Frontiers in Cellular and Infection Microbiology*, 7(October), pp. 1–14. doi: 10.3389/fcimb.2017.00458.

Harrison, X. A. *et al.* (2018) 'A brief introduction to mixed effects modelling and multi-model inference in ecology', *PeerJ*, 6, p. e4794. doi: 10.7717/peerj.4794.

Hirsch, B. T. *et al.* (2013) 'Raccoon Social Networks and the Potential for Disease Transmission', *PLoS ONE*, 8(10), pp. 4–10. doi: 10.1371/journal.pone.0075830.

Hock, K. and Fefferman, N. H. (2012) 'Social organization patterns can lower disease risk without associated disease avoidance or immunity', *Ecological Complexity*. Elsevier B.V., 12, pp. 34–42. doi: 10.1016/j.ecocom.2012.09.003.

Johnson, N. P. A. S. and Mueller, J. (2002) 'Updating the accounts: global mortality of the 1918-1920 "Spanish" influenza pandemic.', *Bulletin of the history of medicine*, 76(1), pp. 105–115.

Jolles, Anna E *et al.* (2008) 'Interactions between Macroparasites and Microparasites Drive Infection Patterns in Free- Ranging African Buffalo Author (s): Anna E . Jolles , Vanessa O . Ezenwa , Rampal S . Etienne , Wendy C . Turner and Han Olff Published by : Wiley Stable URL : http', *Ecology*, 89(8), pp. 2239–2250. Available at: http://onlinelibrary.wiley.com/doi/10.1890/07-0995.1/full.

Jolles, A E *et al.* (2008) 'Interactions between macroparasites and microparasites drive infection patterns in free-ranging African buffalo', *Ecology*, 89(8), pp. 2239–2250. doi: 10.1890/07-0995.1.

Keeling, M. J. *et al.* (2001) 'Dynamics of the 2001 UK foot and mouth epidemic: Stochastic dispersal in a heterogeneous landscape', *Science*, 294(5543), pp. 813–817. doi: 10.1126/science.1065973.

Knowles, S. C. L. *et al.* (2013) 'Stability of within-host-parasite communities in a wild mammal system', *Proceedings of the Royal Society B: Biological Sciences*, 280(1762), p. 20130598. doi: 10.1098/rspb.2013.0598.

Knowles, S. C. L., Fenton, A. and Pedersen, A. B. (2012) 'Epidemiology and fitness effects of wood mouse herpesvirus in a natural host population', *Journal of General Virology*, 93(PART 11), pp. 2447–2456. doi: 10.1099/vir.0.044826-0.

LaDeau, S. L. *et al.* (2008) 'West Nile Virus Revisited: Consequences for North American Ecology', *BioScience*, 58(10), pp. 937–946. doi: 10.1641/b581007.

Leite Dias, J. V. *et al.* (2016) 'Spatial distribution of triatomines in domiciles of an urban area of the Brazilian Southeast region', *Memorias do Instituto Oswaldo Cruz*, 111(1), pp. 43–50. doi: 10.1590/0074-02760150352.

Lello, J. *et al.* (2004) 'Competition and mutulaism among the gut helminths of a mamalian host', *Nature*, 428(April), pp. 20–24. doi: 10.1038/nature02472.1.

Levin, S. A. (1992) 'The Problem of Pattern and Scale in Ecology : The Robert H . MacArthur Award Lecture Stable URL : http://www.jstor.org/stable/1941447 REFERENCES Linked references are available on JSTOR for this article : You may need to log in to JSTOR to access the link', 73(6), pp. 1943–1967.

Madinga, J. *et al.* (2017) 'Geospatial and age-related patterns of Taenia solium taeniasis in the rural health zone of Kimpese, Democratic Republic of Congo', *Acta Tropica*. Elsevier B.V., 165, pp. 100–109. doi: 10.1016/j.actatropica.2016.03.013.

McCallum, H. *et al.* (2017) 'Breaking beta: Deconstructing the parasite transmission function', *Philosophical Transactions of the Royal Society B: Biological Sciences*, 372(1719). doi: 10.1098/rstb.2016.0084.

Nacher, M. *et al.* (2000) 'Ascaris lumbricoides infection is associated with

protection from cerebral malaria', *Parasite Immunology*, 22(March 1999), pp. 107–113. doi: 10.1046/j.1365-3024.2000.00284.x.

Ngowi, H. A. *et al.* (2010) 'Spatial clustering of porcine cysticercosis in Mbulu district, northern Tanzania', *PLoS Neglected Tropical Diseases*, 4(4). doi: 10.1371/journal.pntd.0000652.

Papini, A. and Santosuosso, U. (2017) 'Snow's case revisited: new tool in geographic profiling of epidemiology', *Brazilian Journal of Infectious Diseases*. Elsevier Editora Ltda, 21(1), pp. 112–115. doi: 10.1016/j.bjid.2016.09.010.

Pedersen, A. B. and Antonovics, J. (2013) 'Anthelmintic treatment alters the parasite community in a wild mouse host Anthelmintic treatment alters the parasite community in a wild mouse host', *Biology Letters*, 9(May).

Pedersen, A. B. and Fenton, A. (2007) 'Emphasizing the ecology in parasite community ecology', *Trends in Ecology and Evolution*, 22(3), pp. 133–139. doi: 10.1016/j.tree.2006.11.005.

Petney, T. N. and Andrews, R. H. (1998) 'Multiparasite communities in animals and humans: Frequency, structure and pathogenic significance', *International Journal for Parasitology*, 28(3), pp. 377–393. doi: 10.1016/S0020-7519(97)00189-6.

Picado, A. *et al.* (2011) 'Foot-and-mouth disease in Tanzania from 2001 to 2006', *Transboundary and Emerging Diseases*, 58(1), pp. 44–52. doi: 10.1111/j.1865-1682.2010.01180.x.

Pike, N. (2011) 'Using false discovery rates for multiple comparisons in

ecology and evolution', *Methods in Ecology and Evolution*, 2(3), pp. 278–282. doi: 10.1111/j.2041-210X.2010.00061.x.

Price, S. J. *et al.* (2016) 'Reconstructing the emergence of a lethal infectious disease of wildlife supports a key role for spread through translocations by humans', *Proceedings of the Royal Society B: Biological Sciences*, 283(1839), p. 20160952. doi: 10.1098/rspb.2016.0952.

R Core Team (2018) 'R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria'. Available at: https://www.r-project.org/.

Real, L. A. and Biek, R. (2007) 'Spatial dynamics and genetics of infectious diseases on heterogeneous landscapes', *Journal of the Royal Society Interface*, 4(16), pp. 935–948. doi: 10.1098/rsif.2007.1041.

Ribeiro, F. K. C. *et al.* (2015) 'Genotypic and spatial analysis of mycobacterium tuberculosis transmission in a high-incidence urban setting', *Clinical Infectious Diseases*, 61(5), pp. 758–766. doi: 10.1093/cid/civ365.

Ripley, B. D. (1976) 'The Second-Order Analysis of Stationary Point Processes', *Journal of Applied Probability*, 13(2), pp. 255–266.

Romano, V. *et al.* (2016) 'Modeling infection transmission in primate networks to predict centrality-based risk', *American journal of primatology*, 78(7), pp. 767–779. doi: 10.1002/ajp.22542.

Salathé, M. and Jones, J. H. (2015) 'Dynamics and Control of Diseases in

Networks with Community Structure', *Plos Computational Biology*, 6(4). doi: 10.1371/Citation.

Shiode, N. *et al.* (2015) 'The mortality rates and the space-time patterns of John Snow ' s cholera epidemic map', *International Journal of Health Geographics*. International Journal of Health Geographics, pp. 1–15. doi: 10.1186/s12942-015-0011-y.

Simonsen, L. (1999) 'The global impact of influenza on morbidity and mortality', *Vaccine*, 17, pp. S3–S10.

Smith, D. J. *et al.* (2004) 'Mapping the Antigenic and Genetic Evolution of Influenza Virus', *Science*, 305(July), pp. 371–376. doi: 10.1126/science.1097211.

Smith, D. L. *et al.* (2013) 'Predicting the spatial dynamics of rabies epidemics on heterogeneous landscapes', *Proceedings of the National Academy of Sciences*, 99(6), pp. 3668–3672. doi: 10.1073/pnas.042400799.

Smith, M. J. *et al.* (2009) 'Host-pathogen time series data in wildlife support a transmission function between density and frequency dependence', *Proceedings of the National Academy of Sciences*, 106(19), pp. 7905–7909. doi: 10.1073/pnas.0809145106.

Snow, J. (1855) *On the Mode of Communication of Cholera*. 2nd edn. London: John Churchill, New Burlington Street, England.

Tanner, M. and Michel, A. L. (1999) 'Investigation of the viability of M. bovis under different environmental conditions in the Kruger National

Park', *Onderstepoort journal of veterinary research*, 66, pp. 185–190. Available at: https://pdfs.semanticscholar.org/7457/1d2773f9346850b0ecdcb99ddbf5fd0cf216.pdf.

Telfer, S. *et al.* (2010) 'Species Interactions in a Parasite Community Drive Infection Risk in a Wildlife Population', *Science*, 330, pp. 243–247.

Thrall, P. H., Antonovics, J. and Hall, D. W. (1993) 'Host and pathogen coexistence in vector-borne and venereal diseases characterized by frequency-dependent disease transmission', *American Naturalist*, 142(3), pp. 543–552.

Tobler, W. R. (1970) 'A Computer Movie Simulating Urban Growth in the Detroit Region', *Economic Geography Supplement: Proceedings. International Geographical Union. Commission on Quantitative Methods*, 46, pp. 234–240. Available at: https://www.jstor.org/stable/143141.

Turner, A. K. *et al.* (2014) 'Host-parasite biology in the real world: The field voles of Kielder', *Parasitology*, 141(8), pp. 997–1017. doi: 10.1017/S0031182014000171.

VanderWaal, K. L. *et al.* (2014) 'Linking social and pathogen transmission networks using microbial genetics in giraffe (Giraffa camelopardalis)', *Journal of Animal Ecology*, 83(2), pp. 406–414. doi: 10.1111/1365-2656.12137.

Viiboud, C. *et al.* (2006) 'Hierarchies in the Spread of Influenza', *Science*, 312(April), pp. 447–451.

White, L. A. and Forester, J. D. (2018) 'Dynamic , spatial models of parasite transmission in wildlife : Their structure , applications and remaining challenges', (January 2017), pp. 559–580. doi: 10.1111/1365-2656.12761.

White, L. A., Forester, J. D. and Craft, M. E. (2017) 'Using contact networks to explore mechanisms of parasite transmission in wildlife', *Biological Reviews*, 92(1), pp. 389–409. doi: 10.1111/brv.12236.

Withenshaw, S. M. *et al.* (2016) 'Multihost Bartonella parasites display covert host specificity even when transmitted by generalist vectors', *Journal of Animal Ecology*, 85(6), pp. 1442–1452. doi: 10.1111/1365-2656.12568.

Woolhouse, M. E. J. (2003) 'Foot-and-mouth disease in the UK: What should we do next time?', *Journal of Applied Microbiology*, 94(s1), pp. 126–130. doi: 10.1046/j.1365-2672.94.s1.15.x.

Wu, T.-T. *et al.* (2000) 'Rta of Murine Gammaherpesvirus 68 Reactivates the Complete Lytic Cycle from Latency', *Journal of Virology*, 74(8), pp. 3659–3667. doi: 10.1128/jvi.74.8.3659-3667.2000.

Yakob, L. *et al.* (2013) 'Slaving and release in co-infection control', *Parasites and Vectors*, 6(1), pp. 1–9. doi: 10.1186/1756-3305-6-157.

Yohannan, J. *et al.* (2014) 'Geospatial distribution and clustering of Chlamydia trachomatisin communities undergoing mass azithromycin treatment', *Investigative Ophthalmology and Visual Science*, 55(7), pp. 4144–4150. doi: 10.1167/iovs.14-14148.